

(43) International Publication Date  
3 June 2004 (03.06.2004)

PCT

(10) International Publication Number  
WO 2004/045543 A2(51) International Patent Classification<sup>7</sup>:

A61K

(21) International Application Number:

PCT/US2003/036787

(22) International Filing Date:

14 November 2003 (14.11.2003)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

60/426,137 14 November 2002 (14.11.2002) US  
60/502,050 10 September 2003 (10.09.2003) US

(71) Applicant: DHARMACON, INC. [US/US]; 2650 Crescent Drive, Suite 100, Lafayette, CO 80026 (US).

(72) Inventors: ANASTASIA, Khvorova; 635 Walden #204, Denver, CO (US). ANGELA, Reynolds; 11445 Conifer Ridge Drive, Conifer, CO 80433 (US). DEVIN, Leake; 3050 Krameria Street, Denver, CO 80303 (US). WILLIAM, Marshall; 495 Mohawk Drive, Denver, CO 80303 (US). Stephen, Scaringe; 2746 Prairie Ridge, Lafayette, CO 80026 (US).

(74) Agent: KALOW, David; Kalow &amp; Springut LLP, 488 Madison Ave, 19th Floor, New York, NY 10022 (US).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VC, VN, YU, ZA, ZM, ZW.

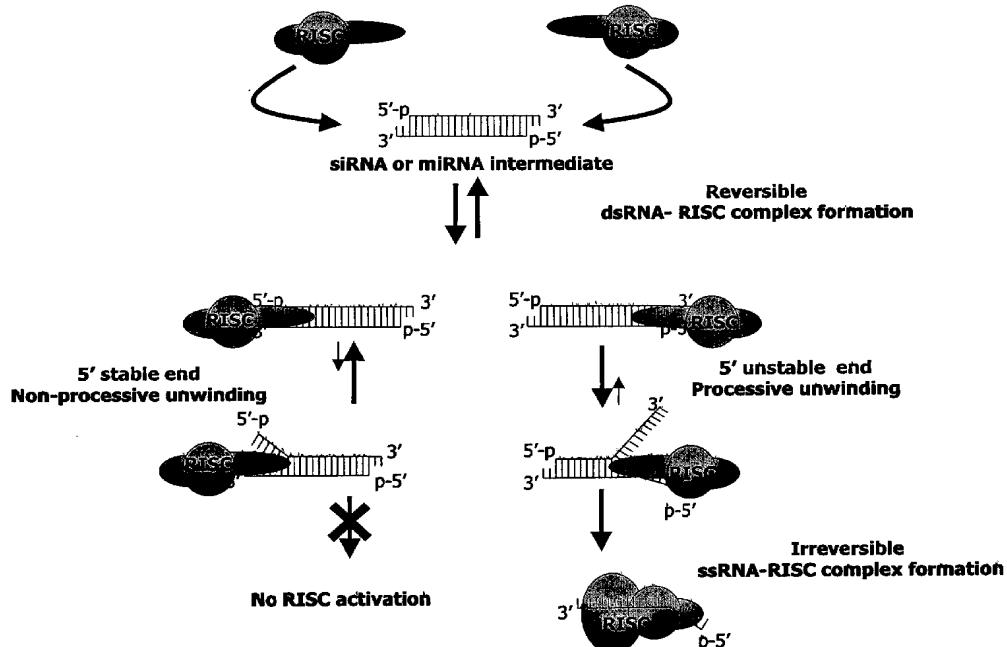
(84) Designated States (regional): ARIPO patent (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

## Published:

— without international search report and to be republished upon receipt of that report

[Continued on next page]

(54) Title: FUNCTIONAL AND HYPERFUNCTIONAL SIRNA



WO 2004/045543 A2

(57) Abstract: Efficient sequence specific gene silencing is possible through the use of siRNA technology. By selecting particular siRNAs by rationale design, one can maximize the generation of an effective gene silencing reagent, as well as methods for silencing genes.



— with sequence listing part of description published separately in electronic form and available upon request from the International Bureau

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

## Functional and Hyperfunctional siRNA

### 5 Cross Reference to Related Applications

This application claims the benefit of the filing date of U.S. Provisional Application Serial No. 60/426,137, filed November 14, 2002, entitled "Combinatorial Pooling Approach for siRNA Induced Gene Silencing and Methods for Selecting siRNA," and U.S. Provisional Application Serial No. 60/502,050, filed September 10, 10 2003, entitled "Methods for Selecting siRNA," the entire disclosures of which are hereby incorporated by reference into the present disclosure.

### Field of Invention

15 The present invention relates to RNA interference ("RNAi").

### Background of the Invention

Relatively recently, researchers observed that double stranded RNA ("dsRNA") could be used to inhibit protein expression. This ability to silence a gene has broad potential for treating human diseases, and many researchers and 20 commercial entities are currently investing considerable resources in developing therapies based on this technology.

25 Double stranded RNA induced gene silencing can occur on at least three different levels: (i) transcription inactivation, which refers to RNA guided DNA or histone methylation; (ii) siRNA induced mRNA degradation; and (iii) mRNA induced transcriptional attenuation.

It is generally considered that the major mechanism of RNA induced silencing (RNA interference, or RNAi) in mammalian cells is mRNA degradation. Initial 30 attempts to use RNAi in mammalian cells focused on the use of long strands of dsRNA. However, these attempts to induce RNAi met with limited success, due in part to the induction of the interferon response, which results in a general, as opposed to a target-specific, inhibition of protein synthesis. Thus, long dsRNA is not a viable option for RNAi in mammalian systems.

More recently it has been shown that when short (18-30 bp) RNA duplexes are introduced into mammalian cells in culture, sequence-specific inhibition of target mRNA can be realized without inducing an interferon response. Certain of these 5 short dsRNAs, referred to as small inhibitory RNAs (“siRNAs”), can act catalytically at sub-molar concentrations to cleave greater than 95% of the target mRNA in the cell. A description of the mechanisms for siRNA activity, as well as some of its applications are described in Provost *et al.*, *Ribonuclease Activity and RNA Binding of Recombinant Human Dicer*, E.M.B.O. J., 2002 Nov. 1; 21(21): 5864 –5874; Tabara *et* 10 *al.*, *The dsRNA Binding Protein RDE-4 Interacts with RDE-1, DCR-1 and a DexH-box Helicase to Direct RNAi in C. elegans*, Cell 2002, June 28;109(7):861-71; Ketting *et al.*, *Dicer Functions in RNA Interference and in Synthesis of Small RNA Involved in Developmental Timing in C. elegans*; Martinez *et al.*, *Single-Stranded Antisense siRNAs Guide Target RNA Cleavage in RNAi*, Cell 2002, Sept. 6; 110(5):563; 15 Hutvagner & Zamore, *A microRNA in a multiple-turnover RNAi enzyme complex*, Science 2002, 297:2056.

From a mechanistic perspective, introduction of long double stranded RNA into plants and invertebrate cells is broken down into siRNA by a Type III 20 endonuclease known as Dicer. Sharp, *RNA interference—2001*, Genes Dev. 2001, 15:485. Dicer, a ribonuclease-III-like enzyme, processes the dsRNA into 19-23 base pair short interfering RNAs with characteristic two base 3’ overhangs. Bernstein, Caudy, Hammond, & Hannon, *Role for a bidentate ribonuclease in the initiation step of RNA interference*, Nature 2001, 409:363. The siRNAs are then incorporated into 25 an RNA-induced silencing complex (RISC) where one or more helicases unwind the siRNA duplex, enabling the complementary antisense strand to guide target recognition. Nykanen, Haley, & Zamore, *ATP requirements and small interfering RNA structure in the RNA interference pathway*, Cell 2001, 107:309. Upon binding to the appropriate target mRNA, one or more endonucleases within the RISC cleaves the 30 target to induce silencing. Elbashir, Lendeckel, & Tuschl, *RNA interference is mediated by 21- and 22-nucleotide RNAs*, Genes Dev 2001, 15:188, **Figure 1**.

The interference effect can be long lasting and may be detectable after many cell divisions. Moreover, RNAi exhibits sequence specificity. Kisielow, M. *et al.*

(2002) *Isoform-specific knockdown and expression of adaptor protein ShcA using small interfering RNA*, J. of Biochemistry 363: 1-5. Thus, the RNAi machinery can specifically knock down one type of transcript, while not affecting closely related mRNA. These properties make siRNA a potentially valuable tool for inhibiting gene 5 expression and studying gene function and drug target validation. Moreover, siRNAs are potentially useful as therapeutic agents against: (1) diseases that are caused by over-expression or misexpression of genes; and (2) diseases brought about by expression of genes that contain mutations.

10       Successful siRNA-dependent gene silencing depends on a number of factors. One of the most contentious issues in RNAi is the question of the necessity of siRNA design, *i.e.*, considering the sequence of the siRNA used. Early work in *C. elegans* and plants circumvented the issue of design by introducing long dsRNA (see, for instance, Fire, A. *et al.* (1998) *Nature* 391:806-811). In this primitive organism, long 15 dsRNA molecules are cleaved into siRNA by Dicer, thus generating a diverse population of duplexes that can potentially cover the entire transcript. While some fraction of these molecules are non-functional (*i.e.* induce little or no silencing) one or more have the potential to be highly functional, thereby silencing the gene of interest and alleviating the need for siRNA design. Unfortunately, due to the interferon 20 response, this same approach is unavailable for mammalian systems. While this effect can be circumvented by bypassing the Dicer cleavage step and directly introducing siRNA, this tactic carries with it the risk that the chosen siRNA sequence may be non-functional or semi-functional.

25       A number of researches have expressed the view that siRNA design is not a crucial element of RNAi. On the other hand, others in the field have begun to explore the possibility that RNAi can be made more efficient by paying attention to the design of the siRNA. Unfortunately, none of the reported methods have provided a satisfactory scheme for reliably selecting siRNA with acceptable levels of 30 functionality. Accordingly, there is a need to develop rational criteria by which to select siRNA with an acceptable level of functionality, and to identify siRNA that have this improved level of functionality, as well as to identify siRNAs that are hyperfunctional.

**Summary of the Invention**

The present invention is directed to increasing the efficiency of RNAi, particularly in mammalian systems. Accordingly, the present invention provides kits, siRNAs and methods for increasing siRNA efficacy.

5

According to one embodiment, the present invention provides a kit for gene silencing, wherein said kit is comprised of a pool of at least two siRNA duplexes, each of which is comprised of a sequence that is complementary to a portion of the sequence of one or more target messenger RNA.

10

According to a second embodiment, the present invention provides a method for optimizing RNA interference by using one or more siRNAs that are optimized according to a formula (or algorithm) selected from:

Formula I

15 Relative functionality of siRNA=  $-(GC/3) + (AU_{15-19}) - (Tm_{20^{\circ}C})^*3 - (G_{13})^*3 - (C_{19})$   
 $+(A_{19})^*2 + (A_3) + (U_{10}) + (A_{14}) - (U_5) - (A_{11})$

Formula II

Relative functionality of siRNA=  $-(GC/3) - (AU_{15-19})^*3 - (G_{13})^*3 - (C_{19}) + (A_{19})^*2$   
20  $+(A_3)$

Formula III

Relative functionality of siRNA=  $-(GC/3) + (AU_{15-19}) - (Tm_{20^{\circ}C})^*3$

25 Formula IV

Relative functionality of siRNA=  $-GC/2 + (AU_{15-19})/2 - (Tm_{20^{\circ}C})^*2 - (G_{13})^*3 - (C_{19}) + (A_{19})^*2 + (A_3) + (U_{10}) + (A_{14}) - (U_5) - (A_{11})$

30 Formula V

Relative functionality of siRNA=  $-(G_{13})^*3 - (C_{19}) + (A_{19})^*2 + (A_3) + (U_{10}) + (A_{14}) - (U_5) - (A_{11})$

Formula VI

Relative functionality of siRNA=-(G<sub>13</sub>)\*3 -(C<sub>19</sub>) +(A<sub>19</sub>)\*2 +(A<sub>3</sub>)

Formula VII

Relative functionality of siRNA=-(GC/2) +(AU<sub>15-19</sub>)/2 -( Tm<sub>20°C</sub>)\*1 -(G<sub>13</sub>)\*3 -(C<sub>19</sub>)

5 +(A<sub>19</sub>)\*3 +(A<sub>3</sub>)\*3 +(U<sub>10</sub>)/2+(A<sub>14</sub>)/2 -(U<sub>5</sub>)/2 -(A<sub>11</sub>)/2

wherein in Formulas I – VII:

Tm<sub>20°C</sub>=1 if the Tm is greater than 20°C;

10 A<sub>19</sub> = 1 if A is the base at position 19 on the sense strand, otherwise its value is 0;

AU<sub>15-19</sub> = 0 – 5 depending on the number of A or U bases on the sense strand at

positions 15 –19;

15 G<sub>13</sub> = 1 if G is the base at position 13 on the sense strand, otherwise its value is 0;

C<sub>19</sub> = 1 if C is the base at position 19 of the sense strand, otherwise its value is 0;

GC= the number of G and C bases in the entire sense strand;

20 A<sub>3</sub>= 1 if A is the base at position 3 on the sense strand, otherwise its value is 0;

A<sub>11</sub>= 1 if A is the base at position 11 on the sense strand, otherwise its value is 0;

A<sub>14</sub>= 1 if A is the base at position 14 on the sense strand, otherwise its value is 0;

25 U<sub>10</sub>= 1 if U is the base at position 10 on the sense strand, otherwise its value is 0;

U<sub>5</sub>= 1 if U is the base at position 5 on the sense strand, otherwise its value is 0; or

30 Formula VIII Relative functionality of siRNA =

(-14)\*G<sub>13</sub>-13\*A<sub>1</sub>-12\*U<sub>7</sub>-11\*U<sub>2</sub>-10\*A<sub>11</sub>-10\*U<sub>4</sub>-10\*C<sub>3</sub>-10\*C<sub>5</sub>-10\*C<sub>6</sub>-  
9\*A<sub>10</sub>-9\*U<sub>9</sub>-9\*C<sub>18</sub>-8\*G<sub>10</sub>-7\*U<sub>1</sub>-7\*U<sub>16</sub>-7\*C<sub>17</sub>-7\*C<sub>19</sub>  
+7\*U<sub>17</sub>+8\*A<sub>2</sub>+8\*A<sub>4</sub>+8\*A<sub>5</sub>+8\*C<sub>4</sub>+9\*G<sub>8</sub>+10\*A<sub>7</sub>+10\*U<sub>18</sub>+11\*A<sub>19</sub>+

11\*C<sub>9</sub>+15\*G<sub>1</sub>+ 18\*A<sub>3</sub>+19\*U<sub>10</sub>-Tm-3\* (GC<sub>total</sub>) - 6\*(GC<sub>15-19</sub>)-  
30\*X; and

Formula IX Relative functionality of siRNA =

5 (14.1)\*A<sub>3</sub>+(14.9)\*A<sub>6</sub>+(17.6)\*A<sub>13</sub>+(24.7)\*A<sub>19</sub>+(14.2)\*U<sub>10</sub>+(10.5)\*  
C<sub>9</sub>+(23.9)\*G<sub>1</sub>+(16.3)\*G<sub>2</sub>+(-12.3)\*A<sub>11</sub>+(-19.3)\*U<sub>1</sub>+(-12.1)\*U<sub>2</sub>+(  
-11)\*U<sub>3</sub>+(-15.2)\*U<sub>15</sub>+(-11.3)\*U<sub>16</sub>+ (-11.8)\*C<sub>3</sub>+(-17.4)\*C<sub>6</sub>+(-  
10.5)\*C<sub>7</sub>+ (-13.7)\*G<sub>13</sub>+(-25.9)\*G<sub>19</sub>-Tm-3\* (GC<sub>total</sub>) - 6\*(GC<sub>15-19</sub>)-  
30\*X

10 wherein

A<sub>1</sub> = 1 if A is the base at position 1 of the sense strand, otherwise its value is 0;  
A<sub>2</sub> = 1 if A is the base at position 2 of the sense strand, otherwise its value is 0;  
A<sub>3</sub> = 1 if A is the base at position 3 of the sense strand, otherwise its value is 0;  
A<sub>4</sub> = 1 if A is the base at position 4 of the sense strand, otherwise its value is 0;  
15 A<sub>5</sub> = 1 if A is the base at position 5 of the sense strand, otherwise its value is 0;  
A<sub>6</sub> = 1 if A is the base at position 6 of the sense strand, otherwise its value is 0;  
A<sub>7</sub> = 1 if A is the base at position 7 of the sense strand, otherwise its value is 0;  
A<sub>10</sub> = 1 if A is the base at position 10 of the sense strand, otherwise its value is 0;  
A<sub>11</sub> = 1 if A is the base at position 11 of the sense strand, otherwise its value is 0;  
20 A<sub>13</sub> = 1 if A is the base at position 13 of the sense strand, otherwise its value is 0;  
A<sub>19</sub> = 1 if A is the base at position 19 of the sense strand, otherwise if another base is  
present or the sense strand is only 18 base pairs in length, its value is 0;

25 C<sub>3</sub> = 1 if C is the base at position 3 of the sense strand, otherwise its value is 0;

C<sub>4</sub> = 1 if C is the base at position 4 of the sense strand, otherwise its value is 0;  
C<sub>5</sub> = 1 if C is the base at position 5 of the sense strand, otherwise its value is 0;  
C<sub>6</sub> = 1 if C is the base at position 6 of the sense strand, otherwise its value is 0;  
C<sub>7</sub> = 1 if C is the base at position 7 of the sense strand, otherwise its value is 0;  
C<sub>9</sub> = 1 if C is the base at position 9 of the sense strand, otherwise its value is 0;  
30 C<sub>17</sub> = 1 if C is the base at position 17 of the sense strand, otherwise its value is 0;  
C<sub>18</sub> = 1 if C is the base at position 18 of the sense strand, otherwise its value is 0;  
C<sub>19</sub> = 1 if C is the base at position 19 of the sense strand, otherwise if another base is  
present or the sense strand is only 18 base pairs in length, its value is 0;

$G_1 = 1$  if G is the base at position 1 on the sense strand, otherwise its value is 0;  
 $G_2 = 1$  if G is the base at position 2 of the sense strand, otherwise its value is 0;  
 $G_8 = 1$  if G is the base at position 8 on the sense strand, otherwise its value is 0;  
 $G_{10} = 1$  if G is the base at position 10 on the sense strand, otherwise its value is 0;  
5       $G_{13} = 1$  if G is the base at position 13 on the sense strand, otherwise its value is 0;  
 $G_{19} = 1$  if G is the base at position 19 of the sense strand, otherwise if another base is  
present or the sense strand is only 18 base pairs in length, its value is 0;

10       $U_1 = 1$  if U is the base at position 1 on the sense strand, otherwise its value is 0;  
 $U_2 = 1$  if U is the base at position 2 on the sense strand, otherwise its value is 0;  
 $U_3 = 1$  if U is the base at position 3 on the sense strand, otherwise its value is 0;  
 $U_4 = 1$  if U is the base at position 4 on the sense strand, otherwise its value is 0;  
 $U_7 = 1$  if U is the base at position 7 on the sense strand, otherwise its value is 0;  
 $U_9 = 1$  if U is the base at position 9 on the sense strand, otherwise its value is 0;  
15       $U_{10} = 1$  if U is the base at position 10 on the sense strand, otherwise its value is 0;  
 $U_{15} = 1$  if U is the base at position 15 on the sense strand, otherwise its value is 0;  
 $U_{16} = 1$  if U is the base at position 16 on the sense strand, otherwise its value is 0;  
 $U_{17} = 1$  if U is the base at position 17 on the sense strand, otherwise its value is 0;  
 $U_{18} = 1$  if U is the base at position 18 on the sense strand, otherwise its value is 0;

20       $GC_{15-19} =$  the number of G and C bases within positions 15 – 19 of the sense strand  
or within positions 15 – 18 if the sense strand is only 18 base pairs in length;  
 $GC_{total} =$  the number of G and C bases in the sense strand;  
25       $Tm = 100$  if the targeting site contains an inverted repeat longer than 4 base pairs,  
otherwise its value is 0; and  
 $X =$  the number of times that the same nucleotide repeats four or more times in a row.

According to a third embodiment, the present invention is directed to a kit  
comprised of at least one siRNA that contains a sequence that is optimized according  
30 to one of the formulas above. Preferably the kit contains at least two optimized  
siRNA, each of which comprises a duplex, wherein one strand of each duplex  
comprises at least eighteen contiguous bases that are complementary to a region of a  
target messenger RNA. For mammalian systems, the siRNA preferably comprises  
between 18 and 30 nucleotide base pairs.

The ability to use the above algorithms, which are not sequence or species specific, allows for the cost-effective selection of optimized siRNAs for specific target sequences. Accordingly, there will be both greater efficiency and reliability in 5 the use of siRNA technologies.

According to a fourth embodiment, the present invention provides a method for developing an siRNA algorithm for selecting functional and hyperfunctional siRNAs for a given sequence. The method comprises:

- 10 (a) selecting a set of siRNAs;
- (b) measuring the gene silencing ability of each siRNA from said set;
- (c) determining the relative functionality of each siRNA;
- 15 (d) determining the amount of improved functionality by the presence or absence of at least one variable selected from the group consisting of the total GC content, melting temperature of the siRNA, GC content at positions 15 –19, the presence or absence of a particular nucleotide at a particular position and the number of times that the same nucleotide repeats within a given sequence; and
- (e) developing an algorithm using the information of step (d).

20

According to this embodiment, preferably the set of siRNAs comprises at least 90 siRNAs from at least one gene, more preferably at least 180 siRNAs from at least two different genes, and most preferably at least 270 and 360 siRNAs from at least three and four different genes, respectively. Additionally, in step (d) the 25 determination is made with preferably at least two, more preferably at least three, even more preferably at least four, and most preferably all of the variables. The resulting algorithm is not target sequence specific.

In a fifth embodiment, the present invention provides rationally designed 30 siRNAs identified using the formulas above.

In a sixth embodiment, the present invention is directed to hyperfunctional siRNA.

For a better understanding of the present invention together with other and further advantages and embodiments, reference is made to the following description taken in conjunction with the examples, the scope of which is set forth in the appended claims.

5

### **Brief Description of the Figures**

**Figure 1** shows a model for siRNA-RISC interactions. RISC has the ability to interact with either end of the siRNA or miRNA molecule. Following binding, the duplex is unwound, and the relevant target is identified, cleaved, and released.

10

**Figure 2** is a representation of the functionality of two hundred and seventy siRNA duplexes that were generated to target human cyclophilin, human diazepam-binding inhibitor (DB), and firefly luciferase.

15

**Figure 3a** is a representation of the silencing effect of 30 siRNAs in three different cell lines, HEK293, DU145, and Hela. **Figure 3b** shows the frequency of different functional groups (>95% silencing (black), >80% silencing (gray), >50% silencing (dark gray), and <50% silencing (white)) based on GC content. In cases where a given bar is absent from a particular GC percentage, no siRNA were identified for that particular group. **Figure 3c** shows the frequency of different functional groups based on melting temperature (Tm). Again, each group has four different divisions: >95% (black), >80% (gray), >50% (dark gray), and <50% (white) silencing.

20

**Figure 4** is a representation of a statistical analysis that revealed correlations between silencing and five sequence-related properties of siRNA: (A) an A at position 19 of the sense strand, (B) an A at position 3 of the sense strand, (C) a U at position 10 of the sense strand, (D) a base other than G at position 13 of the sense strand, and (E) a base other than C at position 19 of the sense strand. All variables were correlated with siRNA silencing of firefly luciferase and human cyclophilin. SiRNAs satisfying the criterion are grouped on the left (Selected) while those that do not, are grouped on the right (Eliminated). Y-axis is "% Silencing of Control." Each position on the X-axis represents a unique siRNA.

**Figures 5 A and 5 B** are representations of firefly luciferase and cyclophilin siRNA panels sorted according to functionality and predicted values using Formula VIII. The siRNA found within the circle represent those that have Formula VIII values (SMARTscores<sup>TM</sup>) above zero. SiRNA outside the indicated area have calculated 5 Formula VIII values that are below zero. Y-axis is “Expression (% Control).” Each position on the X-axis represents a unique siRNA.

**Figure 6A** is a representation of the average internal stability profile (AISP) derived 10 from 270 siRNAs taken from three separate genes (cyclophilin B, DBI and firefly luciferase). Graphs represent AISP values of highly functional, functional, and non-functional siRNA. **Figure 6B** is a comparison between the AISP of naturally derived GFP siRNA (filled squares) and the AISP of siRNA from cyclophilin B, DBI, and luciferase having >90% silencing properties (no fill) for the antisense strand. “DG” is 15 the symbol for  $\Delta G$ , free energy.

**Figure 7** is a histogram showing the differences in duplex functionality upon introduction of basepair mismatches. The X-axis shows the mismatch introduced into 20 the siRNA and the position it is introduced (e.g., 8C->A reveals that position 8 (which normally has a C) has been changed to an A). The Y-axis is “% Silencing (Normalized to Control).”

**Figure 8a** is histogram that shows the effects of 5’ sense and antisense strand modification with 2’-O-methylation on functionality. **Figure 8b** is an expression profile showing a comparison of sense strand off-target effects for IGF1R-3 and 2’-O- 25 methyl IGF1R-3. Sense strand off-targets (lower white box) are not induced when the 5’ end of the sense strand is modified with 2’-O-methyl groups (top white box).

**Figure 9** shows a graph of SMARTscores<sup>TM</sup> versus RNAi silencing values for more than 360 siRNA directed against 30 different genes. SiRNA to the right of the vertical 30 bar represent those siRNA that have desirable SMARTscores<sup>TM</sup>.

**Figures 10A – E** compare the RNAi of five different genes (SEAP, DBI, PLK, Firefly Luciferase, and Renila Luciferase) by varying numbers of randomly selected

siRNA and four rationally designed (SMART-selected) siRNA chosen using the algorithm described in Formula VIII. In addition, RNAi induced by a pool of the four SMART-selected siRNA is reported at two different concentrations (100 and 400nM). 10F is a comparison between a pool of randomly selected EGFR siRNA (Pool 1) and 5 a pool of SMART selected EGFR siRNA (Pool 2). Pool 1, S1—S4 and Pool 2 S1—S4 represent the individual members that made up each respective pool. Note that numbers for random siRNAs represent the position of the 5' end of the sense strand of the duplex. The Y-axis represents the % expression of the control(s). The X-axis is the percent expression of the control.

10

Figure 11 shows the Western blot results from cells treated with siRNA directed against twelve different genes involved in the clathrin-dependent endocytosis pathway (CHC, DynII, CALM, CLCa, CLCb, Eps15, Eps15R, Rab5a, Rab5b, Rab5c,  $\beta$ 2 subunit of AP-2 and EEA.1). SiRNA were selected using Formula VIII. “Pool” 15 represents a mixture of duplexes 1-4. Total concentration of each siRNA in the pool is 25 nM. Total concentration = 4 x 25 = 100 nM.

Figure 12 is a representation of the gene silencing capabilities of rationally-selected siRNA directed against ten different genes (human and mouse cyclophilin, C-myc, 20 human lamin A/C, QB (ubiquinol-cytochrome c reductase core protein I), MEK1 and MEK2, ATE1 (arginyl-tRNA protein transferase), GAPDH, and Eg5). The Y-axis is the percent expression of the control. Numbers 1, 2, 3 and 4 represent individual rationally selected siRNA. “Pool” represents a mixture of the four individual siRNA.

25 Figure 13 is the sequence of the top ten Bcl2 siRNAs as determined by Formula VIII. Sequences are listed 5' to 3'.

Figure 14 is the knockdown by the top ten Bcl2 siRNAs at 100nM concentrations. The Y-axis represents the amount of expression relative to the non-specific (ns) and 30 transfection mixture control.

Figure 15 represents a functional walk where siRNA beginning on every other base pair of a region of the luciferase gene are tested for the ability to silence the luciferase

gene. The Y-axis represents the percent expression relative to a control. The X-axis represents the position of each individual siRNA.

5 **Figure 16** is a histogram demonstrating the inhibition of target gene expression by pools of 2 and 3 siRNAs duplexes taken from the walk described in Figure 15. The Y-axis represents the percent expression relative to control. The X-axis represents the position of the first siRNA in paired pools, or trios of siRNA. For instance, the first paired pool contains siRNA 1 and 3. The second paired pool contains siRNA 3 and 5. Pool 3 (of paired pools) contains siRNA 5 and 7, and so on.

10

**Figure 17** is a histogram demonstrating the inhibition of target gene expression by pools of 4 and 5 siRNA duplexes. The Y-axis represents the percent expression relative to a control. The X-axis represents the positoin of the first siRNA in each pool.

15

**Figure 18** is a histogram demonstrating the inhibition of target gene expression by siRNAs that are ten and twenty basepairs apart. The Y-axis represents the percent expression relative to a control. The X-axis represents the position of the first siRNA in each pool.

20

**Figure 19** shows that pools of siRNAs (dark gray bar) work as well (or better) than the best siRNA in the pool (light gray bar). The Y-axis represents the percent expression relative to a control. The-X axis represents the position of the first siRNA in each pool.

25

**Figure 20** shows that the combination of several semifunctional siRNAs (dark gray) result in a significant improvement of gene expression inhibition over individual (semi-functional; light gray) siRNA. The Y-axis represents the percent expression relative to a control.

30

**Figure 21** shows both pools (Library, Lib) and individual siRNAs in inhibition of gene expression of Beta-Galactosidase, Renilla Luciferase and SEAP (alkaline phosphatase). Numbers on the X-axis indicate the position of the 5'-most nucleotide of the sense strand of the duplex. The Y-axis represents the percent expression of

each gene relative to a control. Libraries contain siRNAs that begin at the following nucleotides: Seap: Lib 1: 206, 766, 812,923, Lib 2: 1117, 1280, 1300, 1487, Lib 3: 206, 766, 812, 923, 1117, 1280, 1300,1487, Lib 4: 206, 812, 1117, 1300, Lib 5: 766, 923, 1280, 1487, Lib 6: 206, 1487; Bgal: Lib 1: 979, 1339, 2029, 2590, Lib 2: 5 1087,1783,2399,3257, Lib 3: 979, 1783, 2590, 3257, Lib 4: 979, 1087, 1339, 1783, 2029, 2399,2590,3257, Lib 5: 979, 1087, 1339, 1783, Lib 6: 2029,2399,2590,3257; Renilla: Lib 1: 174,300,432,568, Lib 2: 592, 633, 729,867, Lib 3: 174, 300, 432, 568, 592, 633,729,867, Lib 4: 174, 432, 592, 729, Lib 5: 300,568,633,867, Lib 6: 592,568.

10

**Figure 22** showS the results of an EGFR and TfnR internalization assay when single gene knockdowns are performed. The Y-axis represents percent internalization relative to control.

15

**Figure 23** shows the results of an EGFR and TfnR internalization assay when multiple genes are knocked down (*e.g.* Rab5a, b, c). The Y-axis represents the percent internalization relative to control.

20

**Figure 24** shows the simultaneous knockdown of four different genes. SiRNAs directed against G6PD, GAPDH, PLK, and UBQ were simultaneously introduced into cells. Twenty-four hours later, cultures were harvested and assayed for mRNA target levels for each of the four genes. A comparison is made between cells transfected with individual siRNAs vs. a pool of siRNAs directed against all four genes.

25

**Figure 25** shows the functionality of ten siRNAs at 0.3nM concentrations.

### **Detailed Description**

#### **Definitions**

30

Unless stated otherwise, the following terms and phrases have the meanings provided below:

siRNA

The term “siRNA” refers to small inhibitory RNA duplexes that induce the RNA interference (RNAi) pathway. These molecules can vary in length (generally between 18-30 basepairS) and contain varying degrees of complementarity to their target mRNA in the antisense strand. Some, but not all, siRNA have unpaired overhanging bases on the 5' or 3' end of the sense strand and/or the antisense strand. The term “siRNA” includes duplexes of two separate strands, as well as single strands that can form hairpin structures comprising a duplex region.

SiRNA may be divided into five (5) groups (non-functional, semi-functional, functional, highly functional, and hyper-functional) based on the level or degree of silencing that they induce in cultured cell lines. As used herein, these definitions are based on a set of conditions where the siRNA is transfected into said cell line at a concentration of 100nM and the level of silencing is tested at a time of roughly 24 hours after transfection, and not exceeding 72 hours after transfection. In this context, “non-functional siRNA” are defined as those siRNA that induce less than 50% (<50%) target silencing. “Semi-functional siRNA” induce 50-79% target silencing. “Functional siRNA” are molecules that induce 80-95% gene silencing. “Highly-functional siRNA” are molecules that induce greater than 95% gene silencing. “Hyperfunctional siRNA” are a special class of molecules. For purposes of this document, hyperfunctional siRNA are defined as those molecules that: (1) induce greater than 95% silencing of a specific target when they are transfected at subnanomolar concentrations (*i.e.*, less than one nanomolar); and/or (2) induce functional (or better) levels of silencing for greater than 96 hours. These relative functionalities (though not intended to be absolutes) may be used to compare siRNAs to a particular target for applications such as functional genomics, target identification and therapeutics.

miRNA

The term “miRNA” refers to microRNA.

Gene silencing

The phrase “gene silencing” refers to a process by which the expression of a specific gene product is lessened or attenuated. Gene silencing can take place by a

variety of pathways. Unless specified otherwise, as used herein, gene silencing refers to decreases in gene product expression that results from RNA interference (RNAi), a defined, though partially characterized pathway whereby small inhibitory RNA (siRNA) act in concert with host proteins (*e.g.* the RNA induced silencing complex, RISC) to degrade messenger RNA (mRNA) in a sequence-dependent fashion. The level of gene silencing can be measured by a variety of means, including, but not limited to, measurement of transcript levels by Northern Blot Analysis, B-DNA techniques, transcription-sensitive reporter constructs, expression profiling (*e.g.* DNA chips), and related technologies. Alternatively, the level of silencing can be measured by assessing the level of the protein encoded by a specific gene. This can be accomplished by performing a number of studies including Western Analysis, measuring the levels of expression of a reporter protein that has *e.g.* fluorescent properties (*e.g.* GFP) or enzymatic activity (*e.g.* alkaline phosphatases), or several other procedures.

15

#### Transfection

The term “transfection” refers to a process by which agents are introduced into a cell. The list of agents that can be transfected is large and includes, but is not limited to, siRNA, sense and/or anti-sense sequences, DNA encoding one or more genes and organized into an expression plasmid, proteins, protein fragments, and more. There are multiple methods for transfecting agents into a cell including, but not limited to, electroporation, calcium phosphate-based transfections, DEAE-dextran-based transfections, lipid-based transfections, molecular conjugate-based transfections (*e.g.* polylysine-DNA conjugates), microinjection and others.

25

#### Target

The term “target” is used in a variety of different forms throughout this document and is defined by the context in which it is used. “Target mRNA” refers to a messenger RNA to which a given siRNA can be directed against. “Target sequence” and “target site” refer to a sequence within the mRNA to which the sense strand of an siRNA shows varying degrees of homology and the antisense strand exhibits varying degrees of complementarity. The term “siRNA target” can refer to the gene, mRNA, or protein against which an siRNA is directed. Similarly “target silencing” can refer to the state of a gene, or the corresponding mRNA or protein.

Off-target silencing and Off-target interference

The phrases “off-target silencing” and “off-target interference” are defined as degradation of mRNA other than the intended target mRNA due to overlapping and/or 5 partial homology with secondary mRNA messages.

SMARTscore™

The term “SMARTscore™” refers to a number determined by applying any of the Formulas I - Formula IX to a given siRNA sequence. The term “SMART-10 selected” or “rationally selected” or “rational selection” refers to siRNA that have been selected on the basis of their SMARTscores™.

Complementary

The term “complementary” refers to the ability of polynucleotides to form 15 base pairs with one another. Base pairs are typically formed by hydrogen bonds between nucleotide units in antiparallel polynucleotide strands. Complementary polynucleotide strands can base pair in the Watson-Crick manner (e.g., A to T, A to U, C to G), or in any other manner that allows for the formation of duplexes. As persons skilled in the art are aware, when using RNA as opposed to DNA, uracil 20 rather than thymine is the base that is considered to be complementary to adenine. However, when a U is denoted in the context of the present invention, the ability to substitute a T is implied, unless otherwise stated.

Perfect complementarity or 100% complementarity refers to the situation in 25 which each nucleotide unit of one polynucleotide strand can hydrogen bond with a nucleotide unit of a second polynucleotide strand. Less than perfect complementarity refers to the situation in which some, but not all, nucleotide units of two strands can hydrogen bond with each other. For example, for two 20-mers, if only two base pairs on each strand can hydrogen bond with each other, the polynucleotide strands exhibit 30 10% complementarity. In the same example, if 18 base pairs on each strand can hydrogen bond with each other, the polynucleotide strands exhibit 90% complementarity. “Substantial complementarity” refers to polynucleotide strands exhibiting 79% or greater complementarity, excluding regions of the polynucleotide strands, such as overhangs, that are selected so as to be noncomplementary.

(“Substantial similarity” refers to polynucleotide strands exhibiting 79% or greater similarity, excluding regions of the polynucleotide strands, such as overhangs, that are selected so as not to be similar.) Thus, for example, two polynucleotides of 29 nucleotide units each, wherein each comprises a di-dT at the 3' terminus such that the 5 duplex region spans 27 bases, and wherein 26 of the 27 bases of the duplex region on each strand are complementary, are substantially complementary since they are 96.3% complementary when excluding the di-dT overhangs.

#### Deoxynucleotide

10 The term “deoxynucleotide” refers to a nucleotide or polynucleotide lacking a hydroxyl group (OH group) at the 2' and/or 3' position of a sugar moiety. Instead, it has a hydrogen bonded to the 2' and/or 3' carbon. Within an RNA molecule that comprises one or more deoxynucleotides, “deoxynucleotide” refers to the lack of an OH group at the 2' position of the sugar moiety, having instead a hydrogen bonded 15 directly to the 2' carbon.

#### Deoxyribonucleotide

20 The terms “deoxyribonucleotide” and “DNA” refer to a nucleotide or polynucleotide comprising at least one sugar moiety that has an H, rather than an OH, at its 2' and/or 3' position.

#### Substantially Similar

The phrase “substantially similar” refers to a similarity of at least 90% with respect to the identity of the bases of the sequence.

25

#### Duplex Region

30 The phrase “duplex region” refers to the region in two complementary or substantially complementary polynucleotides that form base pairs with one another, either by Watson-Crick base pairing or any other manner that allows for a stabilized duplex between polynucleotide strands that are complementary or substantially complementary. For example, a polynucleotide strand having 21 nucleotide units can base pair with another polynucleotide of 21 nucleotide units, yet only 19 bases on each strand are complementary or substantially complementary, such that the “duplex region” has 19 base pairs. The remaining bases may, for example, exist as 5' and 3'

overhangs. Further, within the duplex region, 100% complementarity is not required; substantial complementarity is allowable within a duplex region. Substantial complementarity refers to 79% or greater complementarity. For example, a mismatch in a duplex region consisting of 19 base pairs results in 94.7% complementarity, 5 rendering the duplex region substantially complementary.

### Nucleotide

The term “nucleotide” refers to a ribonucleotide or a deoxyribonucleotide or modified form thereof, as well as an analog thereof. Nucleotides include species that 10 comprise purines, *e.g.*, adenine, hypoxanthine, guanine, and their derivatives and analogs, as well as pyrimidines, *e.g.*, cytosine, uracil, thymine, and their derivatives and analogs.

Nucleotide analogs include nucleotides having modifications in the chemical 15 structure of the base, sugar and/or phosphate, including, but not limited to, 5-position pyrimidine modifications, 8-position purine modifications, modifications at cytosine exocyclic amines, and substitution of 5-bromo-uracil; and 2'-position sugar modifications, including but not limited to, sugar-modified ribonucleotides in which the 2'-OH is replaced by a group such as an H, OR, R, halo, SH, SR, NH<sub>2</sub>, NHR, 20 NR<sub>2</sub>, or CN, wherein R is an alkyl moiety. Nucleotide analogs are also meant to include nucleotides with bases such as inosine, queuosine, xanthine, sugars such as 2'-methyl ribose, non-natural phosphodiester linkages such as methylphosphonates, phosphorothioates and peptides.

25 Modified bases refer to nucleotide bases such as, for example, adenine, guanine, cytosine, thymine, uracil, xanthine, inosine, and queuosine that have been modified by the replacement or addition of one or more atoms or groups. Some examples of types of modifications that can comprise nucleotides that are modified with respect to the base moieties include but are not limited to, alkylated, halogenated, 30 thiolated, aminated, amidated, or acetylated bases, individually or in combination. More specific examples include, for example, 5-propynyluridine, 5-propynylcytidine, 6-methyladenine, 6-methylguanine, N,N-dimethyladenine, 2-propyladenine, 2-propylguanine, 2-aminoadenine, 1-methylinosine, 3-methyluridine, 5-methylcytidine, 5-methyluridine and other nucleotides having a modification at the 5 position, 5-(2-

amino)propyl uridine, 5-halocytidine, 5-halouridine, 4-acetylcytidine, 1-methyladenosine, 2-methyladenosine, 3-methylcytidine, 6-methyluridine, 2-methylguanosine, 7-methylguanosine, 2,2-dimethylguanosine, 5-methylaminoethyluridine, 5-methoxyuridine, deazanucleotides such as 7-deaza-5 adenosine, 6-azouridine, 6-azocytidine, 6-azothymidine, 5-methyl-2-thiouridine, other thio bases such as 2-thiouridine and 4-thiouridine and 2-thiocytidine, dihydrouridine, pseudouridine, queuosine, archaeosine, naphthyl and substituted naphthyl groups, any O- and N-alkylated purines and pyrimidines such as N6-methyladenosine, 5-methylcarbonylmethyluridine, uridine 5-oxyacetic acid, pyridine-4-one, pyridine-2-one, phenyl and modified phenyl groups such as aminophenol or 2,4,6-trimethoxy benzene, modified cytosines that act as G-clamp nucleotides, 8-substituted adenines and guanines, 5-substituted uracils and thymines, azapyrimidines, carboxyhydroxyalkyl nucleotides, carboxyalkylaminoalkyl nucleotides, and alkylcarbonylalkylated nucleotides. Modified nucleotides also include those 10 nucleotides that are modified with respect to the sugar moiety, as well as nucleotides having sugars or analogs thereof that are not ribosyl. For example, the sugar moieties may be, or be based on, mannoses, arabinoses, glucopyranoses, galactopyranoses, 4'-thioribose, and other sugars, heterocycles, or carbocycles.

20 The term nucleotide is also meant to include what are known in the art as universal bases. By way of example, universal bases include but are not limited to 3-nitropyrrole, 5-nitroindole, or nebularine. The term "nucleotide" is also meant to include the N3' to P5' phosphoramidate, resulting from the substitution of a ribosyl 3' oxygen with an amine group.

25 Further, the term nucleotide also includes those species that have a detectable label, such as for example a radioactive or fluorescent moiety, or mass label attached to the nucleotide.

30 Polynucleotide

The term "polynucleotide" refers to polymers of nucleotides, and includes but is not limited to DNA, RNA, DNA/RNA hybrids including polynucleotide chains of regularly and/or irregularly alternating deoxyribosyl moieties and ribosyl moieties (*i.e.*, wherein alternate nucleotide units have an -OH, then and -H, then an -OH, then

an -H, and so on at the 2' position of a sugar moiety), and modifications of these kinds of polynucleotides, wherein the attachment of various entities or moieties to the nucleotide units at any position are included.

5    **Polyribonucleotide**

The term "polyribonucleotide" refers to a polynucleotide comprising two or more modified or unmodified ribonucleotides and/or their analogs. The term "polyribonucleotide" is used interchangeably with the term "oligoribonucleotide."

10    **Ribonucleotide and ribonucleic acid**

The term "ribonucleotide" and the phrase "ribonucleic acid" (RNA), refer to a modified or unmodified nucleotide or polynucleotide comprising at least one ribonucleotide unit. A ribonucleotide unit comprises an hydroxyl group attached to the 2' position of a ribosyl moiety that has a nitrogenous base attached in N-glycosidic linkage at the 1' position of a ribosyl moiety, and a moiety that either allows for linkage to another nucleotide or precludes linkage.

**Detailed Description of the Invention**

20    The present invention is directed to improving the efficiency of gene silencing by siRNA. Through the inclusion of multiple siRNA sequences that are targeted to a particular gene and/or selecting an siRNA sequence based on certain defined criteria, improved efficiency may be achieved.

25    The present invention will now be described in connection with preferred embodiments. These embodiments are presented in order to aid in an understanding of the present invention and are not intended, and should not be construed, to limit the invention in any way. All alternatives, modifications and equivalents that may become apparent to those of ordinary skill upon reading this disclosure are included 30    within the spirit and scope of the present invention.

Furthermore, this disclosure is not a primer on RNA interference. Basic concepts known to persons skilled in the art have not been set forth in detail.

### Optimizing siRNA

According to one embodiment, the present invention provides a method for improving the effectiveness of gene silencing for use to silence a particular gene through the selection of an optimal siRNA. An siRNA selected according to this 5 method may be used individually, or in conjunction with the first embodiment, *i.e.*, with one or more other siRNAs, each of which may or may not be selected by this criteria in order to maximize their efficiency.

The degree to which it is possible to select an siRNA for a given mRNA that 10 maximizes these criteria will depend on the sequence of the mRNA itself. However, the selection criteria will be independent of the target sequence. According to this method, an siRNA is selected for a given gene by using a rational design. That said, rational design can be described in a variety of ways. Rational design is, in simplest terms, the application of a proven set of criteria that enhance the probability of 15 identifying a functional or hyperfunctional siRNA. In one method, rationally designed siRNA can be identified by maximizing one or more of the following criteria:

1. A low GC content, preferably between about 30 –52%.
2. At least 2, preferably at least 3 A or U bases at positions 15- 19 of the 20 siRNA on the sense strand.
3. An A base at position 19 of the sense strand.
4. An A base at position 3 of the sense strand.
5. A U base at position 10 of the sense strand.
6. An A base at position 14 of the sense strand.
7. A base other than C at position 19 of the sense strand.
8. A base other than G at position 13 of the sense strand.
9. A Tm, which refers to the character of the internal repeat that results in 25 inter- or intramolecular structures for one strand of the duplex, that is preferably not stable at greater than 50°C, more preferably not stable at greater than 37°C, even more preferably not stable at greater than 30°C and most preferably not stable at greater than 20°C.
10. A base other than U at position 5 of the sense strand.
11. A base other than A at position 11 of the sense strand.

Criteria 5, 6, 10 and 11 are minor criteria, but are nonetheless desirable. Accordingly, preferably an siRNA will satisfy as many of the aforementioned criteria as possible, more preferably at least 1 – 4 and 7-9, and most preferably all of the 5 criteria

With respect to the criteria, GC content, as well as a high number of AU in positions 15-19, may be important for easement of the unwinding of double stranded siRNA duplex. Duplex unwinding has been shown to be crucial for siRNA 10 functionality *in vivo*.

With respect to criterion 9, the internal structure is measured in terms of the melting temperature of the single strand of siRNA, which is the temperature at which 50% of the molecules will become denatured. With respect to criteria 2 – 8 and 10 – 15 11, the positions refer to sequence positions on the sense strand, which is the strand that is identical to the mRNA.

In one preferred embodiment, at least criteria 1 and 8 are satisfied. In another preferred embodiment, at least criteria 7 and 8 are satisfied. In still another preferred 20 embodiment, at least criteria 1, 8 and 9 are satisfied.

It should be noted that all of the aforementioned criteria regarding sequence position specifics are with respect to the 5' end of the sense strand. Reference is made to the sense strand, because most databases contain information that describes 25 the information of the mRNA. Because according to the present invention a chain can be from 18 to 30 bases in length, and the aforementioned criteria assumes a chain 19 base pairs in length, it is important to keep the aforementioned criteria applicable to the correct bases.

30 When there are only 18 bases, the base pair that is not present is the base pair that is located at the 3' of the sense strand. When there are twenty to thirty bases present, then additional bases are added at the 5' end of the sense chain and occupy positions 1 to 11. Accordingly, with respect to SEQ. ID NO. 0001.  
NNANANNNNUCNAANNNA and SEQ. ID NO. 0028.

GUCNNANANNNUCNAANNNNA, both would have A at position 3, A at position 5, U at position 10, C at position 11, A and position 13, A and position 14 and A at position 19. However, SEQ. ID NO. 0028 would also have C at position -1, U at position -2 and G at position -3.

5

For a 19 base pair siRNA, an optimal sequence of one of the strands may be represented below, where N is any base, A, C, G, or U:

SEQ. ID NO. 0001. NNANANNNUCNAANNNNA  
10 SEQ. ID NO. 0002. NNANANNNUGNAANNNNA  
SEQ. ID NO. 0003. NNANANNNUUNAANNNNA  
SEQ. ID NO. 0004. NNANANNNUCNCANNNA  
SEQ. ID NO. 0005. NNANANNNUGNCANNNA  
SEQ. ID NO. 0006. NNANANNNUUNCANNNA  
15 SEQ. ID NO. 0007. NNANANNNUCNUANNNNA  
SEQ. ID NO. 0008.. NNANANNNUGNUANNNNA  
SEQ. ID NO. 0009. NNANANNNUUNUANNNNA  
SEQ. ID NO. 0010. NNANCNNNUCNAANNNNA  
SEQ. ID NO. 0011. NNANCNNNUGNAANNNNA  
20 SEQ. ID NO. 0012. NNANCNNNUUNAANNNNA  
SEQ. ID NO. 0013. NNANCNNNUCNCANNNA  
SEQ. ID NO. 0014. NNANCNNNUGNCANNNA  
SEQ. ID NO. 0015. NNANCNNNU~~U~~NCANNNA  
SEQ. ID NO. 0016. NANCNNNUCNUANNNNA  
25 SEQ. ID NO. 0017. NNANCNNNUGNUANNNNA  
SEQ. ID NO. 0018. NNANCNNNUUNUANNNNA  
SEQ. ID NO. 0019. NNANGNNNUCNAANNNNA  
SEQ. ID NO. 0020. NNANGNNNUUGNAANNNNA  
SEQ. ID NO. 0021. NNANGNNNUUNAANNNNA  
30 SEQ. ID NO. 0022. NNANGNNNUCNCANNNA  
SEQ. ID NO. 0023. NNANGNNNUUGNCANNNA  
SEQ. ID NO. 0024. NNANGNNNUUNCANNNA  
SEQ. ID NO. 0025. NNANGNNNUCNUANNNNA  
SEQ. ID NO. 0026. NNANGNNNUUGNUANNNNA

SEQ. ID NO. 0027. NNANGNNNNNUUANNNA

In one embodiment, the sequence used as an siRNA is selected by choosing the siRNA that score highest according to one of the following seven algorithms that 5 are represented by Formulas I - VII:

## Formula I

Relative functionality of siRNA=  $-(GC/3) + (AU_{15-19}) - (Tm_{20^{\circ}C})^*3 - (G_{13})^*3 - (C_{19}) + (A_{19})^*2 + (A_3) + (U_{10}) + (A_{14}) - (U_5) - (A_{11})$

10

## Formula II

Relative functionality of siRNA=  $-(GC/3) - (AU_{15-19})^*3 - (G_{13})^*3 - (C_{19}) + (A_{19})^*2 + (A_3)$

15 Formula III

Relative functionality of siRNA=  $-(GC/3) + (AU_{15-19}) - (Tm_{20^{\circ}C})^*3$

## Formula IV

Relative functionality of siRNA=

20  $-GC/2 + (AU_{15-19})/2 - (Tm_{20^{\circ}C})^*2 - (G_{13})^*3 - (C_{19}) + (A_{19})^*2 + (A_3) + (U_{10}) + (A_{14}) - (U_5) - (A_{11})$

## Formula V

Relative functionality of siRNA=  ~~$-(G_{13})^*3 - (C_{19}) + (A_{19})^*2 + (A_3) + (U_{10}) + (A_{14}) - (U_5) - (A_{11})$~~

## Formula VI

Relative functionality of siRNA=  $-(G_{13})^*3 - (C_{19}) + (A_{19})^*2 + (A_3)$

30 Formula VII

Relative functionality of siRNA=  $-(GC/2) + (AU_{15-19})/2 - (Tm_{20^{\circ}C})^*1 - (G_{13})^*3 - (C_{19}) + (A_{19})^*3 + (A_3)^*3 + (U_{10})/2 + (A_{14})/2 - (U_5)/2 - (A_{11})/2$

In Formulas I - VII:

wherein  $A_{19} = 1$  if A is the base at position 19 on the sense strand, otherwise its value is 0,

$AU_{15-19} = 0 - 5$  depending on the number of A or U bases on the sense strand at

5 positions 15 -19;

$G_{13} = 1$  if G is the base at position 13 on the sense strand, otherwise its value is 0;

$C_{19} = 1$  if C is the base at position 19 of the sense strand, otherwise its value is 0;

10  $GC =$  the number of G and C bases in the entire sense strand;

$Tm_{20^{\circ}C} = 1$  if the  $Tm$  is greater than  $20^{\circ}C$ ;

$A_3 = 1$  if A is the base at position 3 on the sense strand, otherwise its value is 0;

$U_{10} = 1$  if U is the base at position 10 on the sense strand, otherwise its value is 0;

15  $A_{14} = 1$  if A is the base at position 14 on the sense strand, otherwise its value is 0;

$U_5 = 1$  if U is the base at position 5 on the sense strand, otherwise its value is 0; and

20  $A_{11} = 1$  if A is the base at position 11 of the sense strand, otherwise its value is 0.

Formulas I –VII provide relative information regarding functionality. When the values for two sequences are compared for a given formula, the relative functionality is ascertained; a higher positive number indicates a greater functionality.

25 For example, in many applications a value of 5 or greater is beneficial.

Additionally, in many applications, more than one of these formulas would provide useful information as to the relative functionality of potential siRNA sequences. However, it is beneficial to have more than one type of formula, because 30 not every formula will be able to help to differentiate among potential siRNA sequences. For example, in particularly high GC mRNAs, formulas that take that parameter into account would not be useful and application of formulas that lack GC elements (e.g., formulas V and VI) might provide greater insights into duplex functionality. Similarly, formula II might be used in situations where hairpin

structures are not observed in duplexes, and formula IV might be applicable for sequences that have higher AU content. Thus, one may consider a particular sequence in light of more than one or even all of these algorithms to obtain the best differentiation among sequences. In some instances, application of a given algorithm 5 may identify an unusually large number of potential siRNA sequences, and in those cases, it may be appropriate to re-analyze that sequence with a second algorithm that is, for instance, more stringent. Alternatively, it is conceivable that analysis of a sequence with a given formula yields no acceptable siRNA sequences (*i.e.* low SMARTscores<sup>TM</sup>). In this instance, it may be appropriate to re-analyze that sequences 10 with a second algorithm that is, for instance, less stringent. In still other instances, analysis of a single sequence with two separate formulas may give rise to conflicting results (*i.e.* one formula generates a set of siRNA with high SMARTscores<sup>TM</sup> while the other formula identifies a set of siRNA with low SMARTscores<sup>TM</sup>). In these instances, it may be necessary to determine which weighted factor(s) (*e.g.* GC 15 content) are contributing to the discrepancy and assessing the sequence to decide whether these factors should or should not be included. Alternatively, the sequence could be analyzed by a third, fourth, or fifth algorithm to identify a set of rationally designed siRNA.

20 The above-referenced criteria are particularly advantageous when used in combination with pooling techniques as depicted in Table I:

**Table I**

| Criteria | Functional Probability |      |      |       |       |       |
|----------|------------------------|------|------|-------|-------|-------|
|          | Oligos                 |      |      | Pools |       |       |
|          | >95%                   | >80% | <70% | >95%  | >80%  | <70%  |
| Current  | 33.0                   | 50.0 | 23.0 | 79.5  | 97.3  | 0.3   |
| New      | 50.0                   | 88.5 | 8.0  | 93.8  | 99.98 | 0.005 |
| (GC)     | 28.0                   | 58.9 | 36.0 | 72.8  | 97.1  | 1.6   |

25 The term “current” refers to Tuschl’s conventional siRNA parameters (Elbashir, S.M. et al. (2002) “Analysis of gene function in somatic mammalian cells using small interfering RNAs” Methods 26: 199-213). “New” refers to the design parameters

described in Formulas I-VII. "GC" refers to criteria that select siRNA solely on the basis of GC content.

As Table I indicates, when more functional siRNA duplexes are chosen, 5 siRNAs that produce <70% silencing drops from 23% to 8% and the number of siRNA duplexes that produce >80% silencing rises from 50% to 88.5%. Further, of the siRNA duplexes with >80% silencing, a larger portion of these siRNAs actually silence >95% of the target expression (the new criteria increases the portion from 33% to 50%). Using this new criteria in pooled siRNAs, shows that, with pooling, the 10 amount of silencing >95% increases from 79.5% to 93.8% and essentially eliminates any siRNA pool from silencing less than 70%.

Table II similarly shows the particularly beneficial results of pooling in combination with the aforementioned criteria. However, Table II, which takes into 15 account each of the aforementioned variables, demonstrates even a greater degree of improvement in functionality.

**Table II**

|            | Functional Probability |         |                |            |         |                |
|------------|------------------------|---------|----------------|------------|---------|----------------|
|            | Oligos                 |         |                | Pools      |         |                |
|            | Functional             | Average | Non-functional | Functional | Average | Non-functional |
| Random     | 20                     | 40      | 50             | 67         | 97      | 3              |
| Criteria 1 | 52                     | 99      | 0.1            | 97         | 93      | 0.0040         |
| Criteria 4 | 89                     | 99      | 0.1            | 99         | 99      | 0.0000         |

The terms "functional," "Average," and "Non-functional" refer to siRNA that exhibit 20 >80%, >50%, and <50% functionality, respectively. Criteria 1 and 4 refer to specific criteria described above.

The above-described algorithms may be used with or without a computer 25 program that allows for the inputting of the sequence of the mRNA and automatically outputs the optimal siRNA. The computer program may, for example, be accessible

from a local terminal or personal computer, over an internal network or over the Internet.

In addition to the formulas above, more detailed algorithms may be used for  
5 selecting siRNA. Preferably, at least one RNA duplex of between 18 and 30 base  
pairs is selected such that it is optimized according a formula selected from:

10 Formula VIII:  $(-14)*G_{13}-13*A_1-12*U_7-11*U_2-10*A_{11}-10*U_4-10*C_3-10*C_5-10*C_6-$   
 $9*A_{10}-9*U_9-9*C_{18}-8*G_{10}-7*U_1-7*U_{16}-7*C_{17}-7*C_{19}$   
 $+7*U_{17}+8*A_2+8*A_4+8*A_5+8*C_4+9*G_8+10*A_7+10*U_{18}+11*A_{19}+$   
 $11*C_9+15*G_1+18*A_3+19*U_{10}-Tm-3*(GC_{total})-6*(GC_{15-19})-$   
30\*X; and

15 Formula IX:  $(14.1)*A_3+(14.9)*A_6+(17.6)*A_{13}+(24.7)*A_{19}+(14.2)*U_{10}+(10.5)*$   
 $C_9+(23.9)*G_1+(16.3)*G_2+(-12.3)*A_{11}+(-19.3)*U_1+(-12.1)*U_2+$   
 $(-11)*U_3+(-15.2)*U_{15}+(-11.3)*U_{16}+(-11.8)*C_3+(-17.4)*C_6+(-$   
 $10.5)*C_7+(-13.7)*G_{13}+(-25.9)*G_{19}-Tm-3*(GC_{total})-6*(GC_{15-19})-$   
30\*X

20

wherein

$A_1 = 1$  if A is the base at position 1 of the sense strand, otherwise its value is 0;  
 $A_2 = 1$  if A is the base at position 2 of the sense strand, otherwise its value is 0;  
 $A_3 = 1$  if A is the base at position 3 of the sense strand, otherwise its value is 0;  
25  $A_4 = 1$  if A is the base at position 4 of the sense strand, otherwise its value is 0;  
 $A_5 = 1$  if A is the base at position 5 of the sense strand, otherwise its value is 0;  
 $A_6 = 1$  if A is the base at position 6 of the sense strand, otherwise its value is 0;  
 $A_7 = 1$  if A is the base at position 7 of the sense strand, otherwise its value is 0;  
 $A_{10} = 1$  if A is the base at position 10 of the sense strand, otherwise its value is 0;  
30  $A_{11} = 1$  if A is the base at position 11 of the sense strand, otherwise its value is 0;  
 $A_{13} = 1$  if A is the base at position 13 of the sense strand, otherwise its value is 0;  
 $A_{19} = 1$  if A is the base at position 19 of the sense strand, otherwise if another base is  
present or the sense strand is only 18 base pairs in length, its value is 0;

$C_3 = 1$  if C is the base at position 3 of the sense strand, otherwise its value is 0;  
 $C_4 = 1$  if C is the base at position 4 of the sense strand, otherwise its value is 0;  
 $C_5 = 1$  if C is the base at position 5 of the sense strand, otherwise its value is 0;  
 $C_6 = 1$  if C is the base at position 6 of the sense strand, otherwise its value is 0;  
5       $C_7 = 1$  if C is the base at position 7 of the sense strand, otherwise its value is 0;  
 $C_9 = 1$  if C is the base at position 9 of the sense strand, otherwise its value is 0;  
 $C_{17} = 1$  if C is the base at position 17 of the sense strand, otherwise its value is 0;  
 $C_{18} = 1$  if C is the base at position 18 of the sense strand, otherwise its value is 0;  
 $C_{19} = 1$  if C is the base at position 19 of the sense strand, otherwise if another base is  
10        present or the sense strand is only 18 base pairs in length, its value is 0;

$G_1 = 1$  if G is the base at position 1 on the sense strand, otherwise its value is 0;  
 $G_2 = 1$  if G is the base at position 2 of the sense strand, otherwise its value is 0;  
 $G_8 = 1$  if G is the base at position 8 on the sense strand, otherwise its value is 0;  
15       $G_{10} = 1$  if G is the base at position 10 on the sense strand, otherwise its value is 0;  
 $G_{13} = 1$  if G is the base at position 13 on the sense strand, otherwise its value is 0;  
 $G_{19} = 1$  if G is the base at position 19 of the sense strand, otherwise if another base is  
      present or the sense strand is only 18 base pairs in length, its value is 0;

20       $U_1 = 1$  if U is the base at position 1 on the sense strand, otherwise its value is 0;  
 $U_2 = 1$  if U is the base at position 2 on the sense strand, otherwise its value is 0;  
 $U_3 = 1$  if U is the base at position 3 on the sense strand, otherwise its value is 0;  
 $U_4 = 1$  if U is the base at position 4 on the sense strand, otherwise its value is 0;  
 $U_7 = 1$  if U is the base at position 7 on the sense strand, otherwise its value is 0;  
25       $U_9 = 1$  if U is the base at position 9 on the sense strand, otherwise its value is 0;  
 $U_{10} = 1$  if U is the base at position 10 on the sense strand, otherwise its value is 0;  
 $U_{15} = 1$  if U is the base at position 15 on the sense strand, otherwise its value is 0;  
 $U_{16} = 1$  if U is the base at position 16 on the sense strand, otherwise its value is 0;  
 $U_{17} = 1$  if U is the base at position 17 on the sense strand, otherwise its value is 0;  
30       $U_{18} = 1$  if U is the base at position 18 on the sense strand, otherwise its value is 0;

$GC_{15-19}$  = the number of G and C bases within positions 15 – 19 of the sense  
strand, or within positions 15 – 18 if the sense strand is only 18 base pairs in  
length;

GC<sub>total</sub> = the number of G and C bases in the sense strand;  
Tm = 100 if the siRNA oligo has the internal repeat longer than 4 base pairs,  
otherwise its value is 0; and  
X = the number of times that the same nucleotide repeats four or more times in a  
5 row.

The above formulas VIII and IX, as well as formulas I – VII, provide methods for selecting siRNA in order to increase the efficiency of gene silencing. A subset of variables of any of the formulas may be used, though when fewer variables are used,  
10 the optimization hierarchy becomes less reliable.

With respect to the variables of the above-referenced formulas, a single letter of A or C or G or U followed by a subscript refers to a binary condition. The binary condition is that either the particular base is present at that particular position  
15 (wherein the value is “1”) or the base is not present (wherein the value is “0”). Because position 19 is optional, *i.e.* there might be only 18 base pairs, when there are only 18 base pairs, any base with a subscript of 19 in the formulas above would have a zero value for that parameter. Before or after each variable is a number followed by \*, which indicates that the value of the variable is to be multiplied or weighed by that  
20 number.

The numbers preceding the variables A, or G, or C, or U in Formulas VIII and IX (or after the variables in Formula I - VII) were determined by comparing the difference in the frequency of individual bases at different positions in functional  
25 siRNA and total siRNA. Specifically, the frequency in which a given base was observed at a particular position in functional groups was compared with the frequency that that same base was observed in the total, randomly selected siRNA set. If the absolute value of the difference between the functional and total values was found to be greater than 6%, that parameter was included in the equation. Thus for  
30 instance, if the frequency of finding a “G” at position 13 (G<sub>13</sub>) is found to be 6% in a given functional group, and the frequency of G<sub>13</sub> in the total population of siRNAs is 20%, the difference between the two values is 6%-20% = -14%. As the absolute value is greater than six (6), this factor (-14) is included in the equation. Thus in Formula VIII, in cases where the siRNA under study has a G in position 13, the accrued value

is  $(-14) * (1) = -14$ . In contrast, when a base other than G is found at position 13, the accrued value is  $(-14) * (0) = 0$ .

When developing a means to optimize siRNAs, the inventors observed that a  
5 bias toward low internal thermodynamic stability of the duplex at the 5'-antisense  
(AS) end is characteristic of naturally occurring miRNA precursors. The inventors  
extended this observation to siRNAs for which functionality had been assessed in  
tissue culture.

10 With respect to the parameter  $GC_{15-19}$ , a value of 0 – 5 will be ascribed  
depending on the number of G or C bases at positions 15 to 19. If there are only 18  
base pairs, the value is between 0 and 4.

15 With respect to the criterion  $GC_{total}$  content, a number from 0 – 30 will be  
ascribed, which correlates to the total number of G and C nucleotides on the sense  
strand, excluding overhangs. Without wishing to be bound by any one theory, it is  
postulated that the significance of the GC content (as well as AU content at positions  
15-19, which is a parameter for formulas III – VII) relates to the easement of the  
unwinding of a double-stranded siRNA duplex. Duplex unwinding is believed to be  
20 crucial for siRNA functionality *in vivo* and overall low internal stability, especially  
low internal stability of the first unwound base pair is believed to be important to  
maintain sufficient processivity of RISC complex-induced duplex unwinding. If the  
duplex has 19 base pairs, those at positions 15-19 on the sense strand will unwind first  
if the molecule exhibits a sufficiently low internal stability at that position. As  
25 persons skilled in the art are aware, RISC is a complex of approximately twelve  
proteins; Dicer is one, but not the only, helicase within this complex. Accordingly,  
although the GC parameters are believed to relate to activity with Dicer, they are also  
important for activity with other RISC proteins.

30 The value of the parameter  $T_m$  is 0 when there are no internal repeats longer  
than (or equal to) four base pairs present in the siRNA duplex; otherwise the value is  
1. Thus for example, if the sequence ACGUACGU, or any other four nucleotide (or  
more) palindrome exists within the structure, the value will be one (1). Alternatively

if the structure ACGGACG, or any other 3 nucleotide (or less) palindrome exists, the value will be zero (0).

The variable "X" refers to the number of times that the same nucleotide occurs 5 contiguously in a stretch of four or more units. If there are, for example, four contiguous As in one part of the sequence and elsewhere in the sequence four contiguous Cs, X =2. Further, if there are two separate contiguous stretches of four of the same nucleotides or eight or more of the same nucleotides in a row, then X=2. However, X does not increase for five, six or seven contiguous nucleotides.

10

Again, when applying Formula VIII or Formula IX to a given mRNA, (the "target RNA" or "target molecule"), one may use a computer program to evaluate the criteria for every sequence of 18 – 30 base pairs or only sequences of a fixed length, e.g., 19 base pairs. Preferably the computer program is designed such that it provides 15 a report ranking of all of the potential siRNAs between 18 and 30 base pairs, ranked according to which sequences generate the highest value. A higher value refers to a more efficient siRNA for a particular target gene. The computer program that may be used, may be developed in any computer language that is known to be useful for scoring nucleotide sequences, or it may be developed with the assistance of 20 commercially available product such as Microsoft's product .net. Additionally, rather than run every sequence through one and/or another formula, one may compare a subset of the sequences, which may be desirable if for example only a subset are available. For instance, it may be desirable to first perform a BLAST (Basic Local Alignment Search Tool) search and to identify sequences that have no homology to 25 other targets. Alternatively, it may be desirable to scan the sequence and to identify regions of moderate GC context, then perform relevant calculations using one of the above-described formulas on these regions. These calculations can be done manually or with the aid of a computer.

30

As with Formulas I – VII, either Formula VIII or Formula IX may be used for a given mRNA target sequence. However, it is possible that according to one or the other formula more than one siRNA will have the same value. Accordingly, it is beneficial to have a second formula by which to differentiate sequences. Formula IX was derived in a similar fashion as Formula VIII, yet used a larger data set and thus

yields sequences with higher statistical correlations to highly functional duplexes. The sequence that has the highest value ascribed to it may be referred to as a “first optimized duplex.” The sequence that has the second highest value ascribed to it may be referred to as a “second optimized duplex.” Similarly, the sequences that have the 5 third and fourth highest values ascribed to them may be referred to as a third optimized duplex and a fourth optimized duplex, respectively. When more than one sequence has the same value, each of them may, for example, be referred to as first optimized duplex sequences or co-first optimized duplexes.

10 SiRNA sequences identified using Formula VIII are contained within the enclosed compact disks. The data included on the enclosed compact disks is described more fully below. The sequences identified by Formula VIII that are disclosed in the compacts disks may be used in gene silencing applications.

15 It should be noted that for Formulas VIII and IX all of the aforementioned criteria are identified as positions on the sense strand when oriented in the 5' to 3' direction as they are identified in connection with Formulas I – VII unless otherwise specified.

20 Formulas I - IX, may be used to select or to evaluate one, or more than one, siRNA in order to optimize silencing. Preferably, at least two optimized siRNAs that have been selected according to at least one of these formulas are used to silence a gene, more preferably at least three and most preferably at least four. The siRNAs may be used individually or together in a pool or kit. Further, they may be applied to 25 a cell simultaneously or separately. Preferably, the at least two siRNAs are applied simultaneously. Pools are particularly beneficial for many research applications. However, for therapeutics, it may be more desirable to employ a single hyperfunctional siRNA as described elsewhere in this application.

30 When planning to conduct gene silencing, and it is necessary to choose between two or more siRNAs, one should do so by comparing the relative values when the siRNA are subjected to one of the formulas above. In general a higher scored siRNA should be used.

Useful applications include, but are not limited to, target validation, gene functional analysis, research and drug discovery, gene therapy and therapeutics. Methods for using siRNA in these applications are well known to persons of skill in the art.

5

Because the ability of siRNA to function is dependent on the sequence of the RNA and not the species into which it is introduced, the present invention is applicable across a broad range of species, including but not limited to all mammalian species, such as humans, dogs, horses, cats, cows, mice, hamsters, chimpanzees and 10 gorillas, as well as other species and organisms such as bacteria, viruses, insects, plants and *C. elegans*.

The present invention is also applicable for use for silencing a broad range of genes, including but not limited to the roughly 45,000 genes of a human genome, and 15 has particular relevance in cases where those genes are associated with diseases such as diabetes, Alzheimer's, cancer, as well as all genes in the genomes of the aforementioned organisms.

The siRNA selected according to the aforementioned criteria or one of the 20 aforementioned algorithms are also, for example, useful in the simultaneous screening and functional analysis of multiple genes and gene families using high throughput strategies, as well as in direct gene suppression or silencing.

#### Development of the Algorithms

25 To identify siRNA sequence features that promote functionality and to quantify the importance of certain currently accepted conventional factors—such as G/C content and target site accessibility—the inventors synthesized an siRNA panel consisting of 270 siRNAs targeting three genes, Human Cyclophilin, Firefly Luciferase, and Human DBI. In all three cases, siRNAs were directed against specific 30 regions of each gene. For Human Cyclophilin and Firefly Luciferase, ninety siRNAs were directed against a 199 bp segment of each respective mRNA. For DBI, 90 siRNAs were directed against a smaller, 109 base pair region of the mRNA. The sequences to which the siRNAs were directed are provided below.

It should be noted that in certain sequences, "t" is present. This is because many databases contain information in this manner. However, the t denotes a uracil residue in mRNA and siRNA. Any algorithm will, unless otherwise specified, process a t in a sequence as a u.

5

Human cyclophilin: 193—390, M60857

SEQ. ID NO. 29:

gttccaaaaacagtgataatttgtggccttagctacaggagagaaggattggctacaaaaacagcaaattccatcgtgt  
aatcaaggacttcatgtatccaggcgaggacttcaccagggagatggcacaggagaaagagcatctacggtgagcg  
10 cttccccatgagaacttcaaactgaagcaactacggcctggctgg

Firefly luciferase: 1434—1631, U47298 (pGL3, Promega)

SEQ. ID NO. 30:

tgaacttcccgccgcccgttgtgtttggagcacggaaagacgatgacggaaaaagagatcgtggattacgtcgccagtca  
15 agtaacaaccgcgaaaaagtgcgcggaggagtgtgtgtggacgaagtaccgaaaggcttaccggaaaactcgacg  
caagaaaaatcagagagatcctcataaaggccaagaagg

DBI, NM\_020548 ( 202-310) (every position)

SEQ. ID NO. 0031:

20 acgggcaaggccaagtggatgcctgaatgagctgaaaggactccaaggaagatccatgaaagcttacatcaaca  
aagtagaagagctaaagaaaaatacggg

A list of the siRNAs appears in Table III (see Examples Section, Example II)

25 The set of duplexes was analyzed to identify correlations between siRNA functionality and other biophysical or thermodynamic properties. When the siRNA panel was analyzed in functional and non-functional subgroups, certain nucleotides were much more abundant at certain positions in functional or non-functional groups. More specifically, the frequency of each nucleotide at each position in highly  
30 functional siRNA duplexes was compared with that of nonfunctional duplexes in order to assess the preference for or against any given nucleotide at every position. These analyses were used to determine important criteria to be included in the siRNA algorithms (Formulas VIII and IX).

The data set was also analyzed for distinguishing biophysical properties of siRNAs in the functional group, such as optimal percent of GC content, propensity for internal structures and regional thermodynamic stability. Of the presented criteria, several are involved in duplex recognition, RISC activation/duplex unwinding, and 5 target cleavage catalysis.

The original data set that was the source of the statistically derived criteria is shown in **Figure 2**. Additionally, this figure shows that random selection yields siRNA duplexes with unpredictable and widely varying silencing potencies as 10 measured in tissue culture using HEK293 cells. In the figure, duplexes are plotted such that each x-axis tick-mark represents an individual siRNA, with each subsequent siRNA differing in target position by two nucleotides for Human Cyclophilin and Firefly Luciferase, and by one nucleotide for Human DBI. Furthermore, the y-axis denotes the level of target expression remaining after transfection of the duplex into 15 cells and subsequent silencing of the target.

SiRNA identified and optimized in this document work equally well in a wide range of cell types. **Figure 3a** shows the evaluation of thirty siRNAs targeting the DBI gene in three cell lines derived from different tissues. Each DBI siRNA displays 20 very similar functionality in HEK293 (ATCC, CRL-1573, human embryonic kidney), HeLa (ATCC, CCL-2, cervical epithelial adenocarcinoma) and DU145 (HTB-81, prostate) cells as determined by the B-DNA assay. Thus, siRNA functionality is determined by the primary sequence of the siRNA and not by the intracellular environment. Additionally, it should be noted that although the present invention 25 provides for a determination of the functionality of siRNA for a given target, the same siRNA may silence more than one gene. For example, the complementary sequence of the silencing siRNA may be present in more than one gene. Accordingly, in these circumstances, it may be desirable not to use the siRNA with highest SMARTscore<sup>TM</sup>. In such circumstances, it may be desirable to use the siRNA with the next highest 30 SMARTscore<sup>TM</sup>.

To determine the relevance of G/C content in siRNA function, the G/C content of each duplex in the panel was calculated and the functional classes of siRNAs (<F50,  $\geq$  F50,  $\geq$  F80,  $\geq$  F95 where F refers to the percent gene silencing) were sorted

accordingly. The majority of the highly-functional siRNAs ( $\geq$ F95) fell within the G/C content range of 36%—52% (**Figure 3B**). Twice as many non-functional (< F50) duplexes fell within the high G/C content groups (>57% GC content) compared to the 36%—52% group. The group with extremely low GC content (26% or less) 5 contained a higher proportion of non-functional siRNAs and no highly-functional siRNAs. The G/C content range of 30%—52% was therefore selected as Criterion I for siRNA functionality, consistent with the observation that a G/C range 30%—70% promotes efficient RNAi targeting. Application of this criterion alone provided only a marginal increase in the probability of selecting functional siRNAs from the panel: 10 selection of F50 and F95 siRNAs was improved by 3.6% and 2.2%, respectively. The siRNA panel presented here permitted a more systematic analysis and quantification of the importance of this criterion than that used previously.

A relative measure of local internal stability is the A/U base pair (bp) content; 15 therefore, the frequency of A/U bp was determined for each of the five terminal positions of the duplex (5' sense (S)/5' antisense (AS)) of all siRNAs in the panel. Duplexes were then categorized by the number of A/U bp in positions 1—5 and 15—19 of the sense strand. The thermodynamic flexibility of the duplex 5'-end (positions 1—5; S) did not appear to correlate appreciably with silencing potency, while that of 20 the 3'-end (positions 15—19; S) correlated with efficient silencing. No duplexes lacking A/U bp in positions 15—19 were functional. The presence of one A/U bp in this region conferred some degree of functionality, but the presence of three or more A/Us was preferable and therefore defined as Criterion II. When applied to the test panel, only a marginal increase in the probability of functional siRNA selection was 25 achieved: a 1.8% and 2.3% increase for F50 and F95 duplexes, respectively (**Table IV**).

The complementary strands of siRNAs that contain internal repeats or 30 palindromes may form internal fold-back structures. These hairpin-like structures exist in equilibrium with the duplexed form effectively reducing the concentration of functional duplexes. The propensity to form internal hairpins and their relative stability can be estimated by predicted melting temperatures. High  $T_m$  reflects a tendency to form hairpin structures. Lower  $T_m$  values indicate a lesser tendency to form hairpins. When the functional classes of siRNAs were sorted by  $T_m$  (**Figure 3c**),

the following trends were identified: duplexes lacking stable internal repeats were the most potent silencers (no F95 duplex with predicted hairpin structure  $T_m > 60$  °C). In contrast, about 60% of the duplexes in the groups having internal hairpins with calculated  $T_m$  values less than 20 °C were F80. Thus, the stability of internal repeats 5 is inversely proportional to the silencing effect and defines Criterion III (predicted hairpin structure  $T_m \leq 20$  °C).

#### Sequence-based determinants of siRNA functionality

When the siRNA panel was sorted into functional and non-functional groups, 10 the frequency of a specific nucleotide at each position in a functional siRNA duplex was compared with that of a nonfunctional duplex in order to assess the preference for or against a certain nucleotide. **Figure 4** shows the results of these queries and the subsequent resorting of the data set (from Figure 2). The data is separated into two sets: those duplexes that meet the criteria, a specific nucleotide in a certain position - 15 grouped on the left (Selected) and those that do not - grouped on the right (Eliminated). The duplexes are further sorted from most functional to least functional with the y-axis of Figure 4a-e representing the % expression *i.e.* the amount of silencing that is elicited by the duplex (Note: each position on the X-axis represents a different duplex). Statistical analysis revealed correlations between silencing and 20 several sequence-related properties of siRNAs. **Figure 4** and **Table IV** show quantitative analysis for the following five sequence-related properties of siRNA: (A) an A at position 19 of the sense strand; (B) an A at position 3 of the sense strand; (C) a U at position 10 of the sense strand; (D) a base other than G at position 13 of the sense strand; and (E) a base other than C at position 19 of the sense strand.

25

When the siRNAs in the panel were evaluated for the presence of an A at position 19 of the sense strand, the percentage of non-functional duplexes decreased from 20% to 11.8%, and the percentage of F95 duplexes increased from 21.7% to 29.4% (**Table IV**). Thus, the presence of an A in this position defined Criterion IV.

30

Another sequence-related property correlated with silencing was the presence of an A in position 3 of the sense strand (**Figure 4b**). Of the siRNAs with A3, 34.4% were F95, compared with 21.7% randomly selected siRNAs. The presence of a U base in position 10 of the sense strand exhibited an even greater impact (**Figure 4c**).

Of the duplexes in this group, 41.7% were F95. These properties became criteria V and VI, respectively.

Two negative sequence-related criteria that were identified also appear on 5 **Figure 4**. The absence of a G at position 13 of the sense strand, conferred a marginal increase in selecting functional duplexes (**Figure 4d**). Similarly, lack of a C at position 19 of the sense strand also correlated with functionality (**Figure 4e**). Thus, among functional duplexes, position 19 was most likely occupied by A, and rarely occupied by C. These rules were defined as criteria VII and VIII, respectively.

10

Application of each criterion individually provided marginal but statistically significant increases in the probability of selecting a potent siRNA. Although the results were informative, the inventors sought to maximize potency and therefore consider multiple criteria or parameters. Optimization is particularly important when 15 developing therapeutics. Interestingly, the probability of selecting a functional siRNA based on each thermodynamic criteria was 2%—4% higher than random, but 4%—8% higher for the sequence-related determinates. Presumably, these sequence-related increases reflect the complexity of the RNAi mechanism and the multitude of protein-RNA interactions that are involved in RNAi-mediated silencing.

20

Table IV

| <b>Criterion</b>  | <b>% Functional</b> |       | <b>Improvement over Random</b> |
|---|---------------------|-------|--------------------------------|
| 5<br>I. 30%—52% G/C content   | < F50               | 16.4% | -3.6%                          |
|   | ≥ F50               | 83.6% | 3.6%                           |
|   | ≥ F80               | 60.4% | 4.3%                           |
|   | ≥ F95               | 23.9% | 2.2%                           |
| 10<br>II. At least 3 A/U bases at positions 15—19 of the sense strand                                       | < F50               | 18.2% | -1.8%                          |
|   | ≥ F50               | 81.8% | 1.8%                           |
|   | ≥ F80               | 59.7% | 3.6%                           |
|   | ≥ F95               | 24.0% | 2.3%                           |
| 15<br>III. Absence of internal repeats, as measured by $T_m$ of secondary structure $\leq 20^\circ\text{C}$ | < F50               | 16.7% | -3.3%                          |
|   | ≥ F50               | 83.3% | 3.3%                           |
|   | ≥ F80               | 61.1% | 5.0%                           |
|   | ≥ F95               | 24.6% | 2.9%                           |
| 20<br>IV. An A base at position 19 of the sense strand  | < F50               | 11.8% | -8.2%                          |
|   | ≥ F50               | 88.2% | 8.2%                           |
|   | ≥ F80               | 75.0% | 18.9%                          |
|   | ≥ F95               | 29.4% | 7.7%                           |
| 25<br>V. An A base at position 3 of the sense strand  | < F50               | 17.2% | -2.8%                          |
|   | ≥ F50               | 82.8% | 2.8%                           |
|   | ≥ F80               | 62.5% | 6.4%                           |
|   | ≥ F95               | 34.4% | 12.7%                          |
| 20<br>VI. A U base at position 10 of the sense strand   | < F50               | 13.9% | -6.1%                          |
|   | ≥ F50               | 86.1% | 6.1%                           |
|   | ≥ F80               | 69.4% | 13.3%                          |
|   | ≥ F95               | 41.7% | 20%                            |
| 25<br>VII. A base other than C at position 19 of the sense strand   | < F50               | 18.8% | -1.2%                          |
|   | ≥ F50               | 81.2% | 1.2%                           |
|   | ≥ F80               | 59.7% | 3.6%                           |
|   | ≥ F95               | 24.2% | 2.5%                           |
| 30<br>VIII. A base other than G at position 13 of the sense strand  | < F50               | 15.2% | -4.8%                          |
|   | ≥ F50               | 84.8% | 4.8%                           |
|   | ≥ F80               | 61.4% | 5.3%                           |
|   | ≥ F95               | 26.5% | 4.8%                           |

25

#### The siRNA selection algorithm

In an effort to improve selection further, all identified criteria, including but not limited to those listed in Table IV were combined into the algorithms embodied in Formula VIII and Formula IX. Each siRNA was then assigned a score (referred to as 30 a SMARTscore™) according to the values derived from the formulas. Duplexes that scored higher than 0 or 20, for Formulas VIII and IX, respectively, effectively selected a set of functional siRNAs and excluded all non-functional siRNAs. Conversely, all duplexes scoring lower than 0 and 20 according to formulas VIII and

IX, respectively, contained some functional siRNAs but included all non-functional siRNAs. A graphical representation of this selection is shown in **Figure 5**.

The methods for obtaining the seven criteria embodied in Table IV are  
5 illustrative of the results of the process used to develop the information for Formulas VIII and IX. Thus similar techniques were used to establish the other variables and their multipliers. As described above, basic statistical methods were used to determine the relative values for these multipliers.

10 To determine the value for "Improvement over Random" the difference in the frequency of a given attribute (*e.g.* GC content, base preference) at a particular position is determined between individual functional groups (*e.g.* <F50) and the total siRNA population studied (*e.g.* 270 siRNA molecules selected randomly). Thus, for instance, in Criterion I (30%-52% GC content) members of the <F50 group were  
15 observed to have GC contents between 30-52% in 16.4% of the cases. In contrast, the total group of 270 siRNAs had GC contents in this range, 20% of the time. Thus for this particular attribute, there is a small negative correlation between 30%-52% GC content and this functional group (*i.e.* 16.4%-20% = -3.6%). Similarly, for Criterion VI, (a "U" at position 10 of the sense strand), the >F95 group contained a "U" at this  
20 position 41.7% of the time. In contrast, the total group of 270 siRNAs had a "U" at this position 21.7% of the time, thus the improvement over random is calculated to be 20% (or 41.7%-21.7%).

#### Identifying The Average Internal Stability Profile of Strong siRNA

25 In order to identify an internal stability profile that is characteristic of strong siRNA, 270 different siRNAs derived from the cyclophilin B, the diazepam binding inhibitor (DBI), and the luciferase gene were individually transfected into HEK293 cells and tested for their ability to induce RNAi of the respective gene. Based on their performance in the *in vivo* assay, the sequences were then subdivided into three  
30 groups, (i) >95% silencing; (ii) 80-95% silencing; and (iii) less than 50% silencing. Sequences exhibiting 51-84% silencing were eliminated from further consideration to reduce the difficulties in identifying relevant thermodynamic patterns.

Following the division of siRNA into three groups, a statistical analysis was performed on each member of each group to determine the average internal stability profile (AISP) of the siRNA. To accomplish this the Oligo 5.0 Primer Analysis Software and other related statistical packages (e.g. Excel) were exploited to

5 determine the internal stability of pentamers using the nearest neighbor method described by Freier *et al.*, (1986) *Improved free-energy parameters for predictions of RNA duplex stability*, Proc Natl. Acad. Sci. U. S. A. 83(24): 9373-7. Values for each group at each position were then averaged, and the resulting data were graphed on a linear coordinate system with the Y-axis expressing the  $\Delta G$  (free energy) values in

10 kcal/mole and the X-axis identifying the position of the base relative to the 5' end.

The results of the analysis identified multiple key regions in siRNA molecules that were critical for successful gene silencing. At the 3'-most end of the sense strand (5' antisense), highly functional siRNA (>95% gene silencing, see **Figure 6a**, >F95)

15 have a low internal stability (AISP of position 19 = ~ -7.6kcal/mol). In contrast low-efficiency siRNA (*i.e.* those exhibiting less than 50% silencing, <F50) display a distinctly different profile, having high  $\Delta G$  values (~ -8.4kcal/mol) for the same position. Moving in a 5' (sense strand) direction, the internal stability of highly efficient siRNA rises (position 12 = ~ -8.3kcal/mol) and then drops again (position 7

20 = ~ -7.7kcal/mol) before leveling off at a value of approximately -8.1kcal/mol for the 5' terminus. SiRNA with poor silencing capabilities show a distinctly different profile. While the AISP value at position 12 is nearly identical with that of strong siRNAs, the values at positions 7 and 8 rise considerably, peaking at a high of ~ -9.0 kcal/mol. In addition, at the 5' end of the molecule the AISP profile of strong and

25 weak siRNA differ dramatically. Unlike the relatively strong values exhibited by siRNA in the >95% silencing group, siRNAs that exhibit poor silencing activity have weak AISP values (-7.6, -7.5, and -7.5 kcal/mol for positions 1, 2 and 3 respectively).

Overall the profiles of both strong and weak siRNAs form distinct sinusoidal shapes that are roughly 180° out-of-phase with each other. While these thermodynamic descriptions define the archetypal profile of a strong siRNA, it will likely be the case that neither the  $\Delta G$  values given for key positions in the profile or the absolute position of the profile along the Y-axis (*i.e.* the  $\Delta G$  -axis) are absolutes.

Profiles that are shifted upward or downward (*i.e.* having on an average, higher or lower values at every position) but retain the relative shape and position of the profile along the X-axis can be foreseen as being equally effective as the model profile described here. Moreover, it is likely that siRNA that have strong or even stronger 5 gene-specific silencing effects might have exaggerated  $\Delta G$  values (either higher or lower) at key positions. Thus, for instance, it is possible that the 5'-most position of the sense strand (position 19) could have  $\Delta G$  values of 7.4 kcal/mol or lower and still be a strong siRNA if, for instance, a G-C  $\rightarrow$  G-T/U mismatch were substituted at position 19 and altered duplex stability. Similarly, position 12 and position 7 could 10 have values above 8.3 kcal/mol and below 7.7 kcal/mole, respectively, without abating the silencing effectiveness of the molecule. Thus, for instance, at position 12, a stabilizing chemical modification (*e.g.* a chemical modification of the 2' position of the sugar backbone) could be added that increases the average internal stability at that position. Similarly, at position 7, mismatches similar to those described previously 15 could be introduced that would lower the  $\Delta G$  values at that position.

Lastly, it is important to note that while functional and non-functional siRNA were originally defined as those molecules having specific silencing properties, both broader or more limiting parameters can be used to define these molecules. As used 20 herein, unless otherwise specified, “non-functional siRNA” are defined as those siRNA that induce less than 50% (<50%) target silencing, “semi-functional siRNA” induce 50-79% target silencing, “functional siRNA” are molecules that induce 80-95% gene silencing, and “highly-functional siRNA” are molecules that induce great than 95% gene silencing. These definitions are not intended to be rigid and can vary 25 depending upon the design and needs of the application. For instance, it is possible that a researcher attempting to map a gene to a chromosome using a functional assay, may identify an siRNA that reduces gene activity by only 30%. While this level of gene silencing may be “non-functional” for *e.g.* therapeutic needs, it is sufficient for gene mapping purposes and is, under these uses and conditions, “functional.” For 30 these reasons, functional siRNA can be defined as those molecules having greater than 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90% silencing capabilities at 100nM transfection conditions. Similarly, depending upon the needs of the study and/or application, non-functional and semi-functional siRNA can be defined as

having different parameters. For instance, semi-functional siRNA can be defined as being those molecules that induce 20%, 30%, 40%, 50%, 60%, or 70% silencing at 100nM transfection conditions. Similarly, non-functional siRNA can be defined as being those molecules that silence gene expression by less than 70%, 60%, 50%, 5 40%, 30%, or less. Nonetheless, unless otherwise stated, the descriptions stated in the “Definitions” section of this text should be applied.

Functional attributes can be assigned to each of the key positions in the AISPs of strong siRNA. The low 5' (sense strand) AISPs values of strong siRNAs may be 10 necessary for determining which end of the molecule enters the RISC complex. In contrast, the high and low AISPs values observed in the central regions of the molecule may be critical for siRNA-target mRNA interactions and product release, respectively.

15 If the AISPs values described above accurately define the thermodynamic parameters of strong siRNA, it would be expected that similar patterns would be observed in strong siRNA isolated from nature. Natural siRNAs exist in a harsh, RNase-rich environment and it can be hypothesized that only those siRNA that exhibit heightened affinity for RISC (*i.e.* siRNA that exhibit an average internal 20 stability profile similar to those observed in strong siRNA) would survive in an intracellular environment. This hypothesis was tested using GFP-specific siRNA isolated from *N. benthamiana*. Llave *et al.* (2002) *Endogenous and Silencing-Associated Small RNAs in Plants*, *The Plant Cell* 14, 1605-1619, introduced long double-stranded GFP-encoding RNA into plants and subsequently re-isolated GFP- 25 specific siRNA from the tissues. The AISPs of fifty-nine of these GFP-siRNA were determined, averaged, and subsequently plotted alongside the AISPs profile obtained from the cyclophilin B/DBI/ luciferase siRNA having >90% silencing properties (**Figure 6b**). Comparison of the two groups show that profiles are nearly identical. This finding validates the information provided by the internal stability profiles and 30 demonstrates that: (1) the profile identified by analysis of the cyclophilin B/DBI/ luciferase siRNAs are not gene specific; and (2) AISPs values can be used to search for strong siRNAs in a variety of species.

Both chemical modifications and base-pair mismatches can be incorporated into siRNA to alter the duplex's AISP and functionality. For instance, introduction of mismatches at positions 1 or 2 of the sense strand destabilized the 5' end of the sense strand and increases the functionality of the molecule (see Luc, **Figure 7**). Similarly, 5 addition of 2'-O-methyl groups to positions 1 and 2 of the sense strand can also alter the AISP and (as a result) increase both the functionality of the molecule and eliminate off-target effects that results from sense strand homology with the unrelated targets (**Figures 8a, 8b**).

10 **Rationale for Criteria in a Biological Context**

The fate of siRNA in the RNAi pathway may be described in 5 major steps: (1) duplex recognition and pre-RISC complex formation; (2) ATP-dependent duplex unwinding/strand selection and RISC activation; (3) mRNA target identification; (4) mRNA cleavage, and (5) product release (Figure 1). Given the level of nucleic acid-protein interactions at each step, siRNA functionality is likely influenced by specific 15 biophysical and molecular properties that promote efficient interactions within the context of the multi-component complexes. Indeed, the systematic analysis of the siRNA test set identified multiple factors that correlate well with functionality. When combined into a single algorithm, they proved to be very effective in selecting active 20 siRNAs.

The factors described here may also be predictive of key functional associations important for each step in RNAi. For example, the potential formation of internal hairpin structures correlated negatively with siRNA functionality. 25 Complementary strands with stable internal repeats are more likely to exist as stable hairpins thus decreasing the effective concentration of the functional duplex form. This suggests that the duplex is the preferred conformation for initial pre-RISC association. Indeed, although single complementary strands can induce gene silencing, the effective concentration required is at least two orders of magnitude 30 higher than that of the duplex form.

siRNA-pre-RISC complex formation is followed by an ATP-dependent duplex unwinding step and "activation" of the RISC. The siRNA functionality was shown to correlate with overall low internal stability of the duplex and low internal stability of

the 3' sense end (or differential internal stability of the 3' sense compare to the 5' sense strand), which may reflect strand selection and entry into the RISC. Overall duplex stability and low internal stability at the 3' end of the sense strand were also correlated with siRNA functionality. Interestingly, siRNAs with very high and very 5 low overall stability profiles correlate strongly with non-functional duplexes. One interpretation is that high internal stability prevents efficient unwinding while very low stability reduces siRNA target affinity and subsequent mRNA cleavage by the RISC.

10 Several criteria describe base preferences at specific positions of the sense strand and are even more intriguing when considering their potential mechanistic roles in target recognition and mRNA cleavage. Base preferences for A at position 19 of the sense strand but not C, are particularly interesting because they reflect the same base preferences observed for naturally occurring miRNA precursors. That is, among 15 the reported miRNA precursor sequences 75% contain a U at position 1 which corresponds to an A in position 19 of the sense strand of siRNAs, while G was under-represented in this same position for miRNA precursors. These observations support the hypothesis that both miRNA precursors and siRNA duplexes are processed by very similar if not identical protein machinery. The functional interpretation of the 20 predominance of a U/A base pair is that it promotes flexibility at the 5' antisense ends of both siRNA duplexes and miRNA precursors and facilitates efficient unwinding and selective strand entrance into an activated RISC.

25 Among the criteria associated with base preferences that are likely to influence mRNA cleavage or possibly product release, the preference for U at position 10 of the sense strand exhibited the greatest impact, enhancing the probability of selecting an F80 sequence by 13.3%. Activated RISC preferentially cleaves target mRNA between nucleotides 10 and 11 relative to the 5' end of the complementary targeting strand. Therefore, it may be that U, the preferred base for 30 most endoribonucleases, at this position supports more efficient cleavage. Alternatively, a U/A bp between the targeting siRNA strand and its cognate target mRNA may create an optimal conformation for the RISC-associated "slicing" activity.

### Pooling

According to another embodiment, the present invention provides a pool of at least two siRNAs, preferably in the form of a kit or therapeutic reagent, wherein one strand of each of the siRNAs, the sense strand comprises a sequence that is substantially similar to a sequence within a target mRNA. The opposite strand, the antisense strand, will preferably comprise a sequence that is substantially complementary to that of the target mRNA. More preferably, one strand of each siRNA will comprise a sequence that is identical to a sequence that is contained in the target mRNA. Most preferably, each siRNA will be 19 base pairs in length, and one strand of each of the siRNAs will be 100% complementary to a portion of the target mRNA.

By increasing the number of siRNAs directed to a particular target using a pool or kit, one is able both to increase the likelihood that at least one siRNA with satisfactory functionality will be included, as well as to benefit from additive or synergistic effects. Further, when two or more siRNAs directed against a single gene do not have satisfactory levels of functionality alone, if combined, they may satisfactorily promote degradation of the target messenger RNA and successfully inhibit translation. By including multiple siRNAs in the system, not only is the probability of silencing increased, but the economics of operation are also improved when compared to adding different siRNAs sequentially. This effect is contrary to the conventional wisdom that the concurrent use of multiple siRNA will negatively impact gene silencing (e.g. Holen, T. *et al.* (2003) "Similar behavior of single strand and double strand siRNAs suggests they act through a common RNAi pathway." NAR 31: 2401-21407).

In fact, when two siRNAs were pooled together, 54% of the pools of two siRNAs induced more than 95% gene silencing. Thus, a 2.5-fold increase in the percentage of functionality was achieved by randomly combining two siRNAs. Further, over 84% of pools containing two siRNAs induced more than 80% gene silencing.

More preferably, the kit is comprised of at least three siRNAs, wherein one strand of each siRNA comprises a sequence that is substantially similar to a sequence

of the target mRNA and the other strand comprises a sequence that is substantially complementary to the region of the target mRNA. As with the kit that comprises at least two siRNAs, more preferably one strand will comprise a sequence that is identical to a sequence that is contained in the mRNA and another strand that is 100% 5 complementary to a sequence that is contained in the mRNA. During experiments, when three siRNAs were combined together, 60% of the pools induced more than 95% gene silencing and 92% of the pools induced more than 80% gene silencing.

Further, even more preferably, the kit is comprised of at least four siRNAs, 10 wherein one strand of each siRNA comprises a sequence that is substantially similar to a region of the sequence of the target mRNA, and the other strand comprises a sequence that is substantially complementary to the region of the target mRNA. As with the kit or pool that comprises at least two siRNAs, more preferably one strand of each of the siRNA duplexes will comprise a sequence that is identical to a sequence 15 that is contained in the mRNA, and another strand that is 100% complementary to a sequence that is contained in the mRNA.

Additionally, kits and pools with at least five, at least six, and at least seven 20 siRNAs may also be useful with the present invention. For example, pools of five siRNA induced 95% gene silencing with 77% probability and 80% silencing with 98.8% probability. Thus, pooling of siRNAs together can result in the creation of a target-specific silencing reagent with almost a 99% probability of being functional. The fact that such high levels of success are achievable using such pools of siRNA, 25 enables one to dispense with costly and time-consuming target-specific validation procedures.

For this embodiment, as well as the other aforementioned embodiments, each 30 of the siRNAs within a pool will preferably comprise between 18 and 30 base pairs, more preferably between 18 and 25 base pairs, and most preferably 19 base pairs. Within each siRNA, preferably at least 18 contiguous bases of the antisense strand will be 100% complementary to the target mRNA. More preferably, at least 19 contiguous bases of the antisense strand will be 100% complementary to the target mRNA. Additionally, there may be overhangs on either the sense strand or the antisense strand, and these overhangs may be at either the 5' end or the 3' end of

either of the strands, for example there may be one or more overhangs of 1-6 bases. When overhangs are present, they are not included in the calculation of the number of base pairs. The two nucleotide 3' overhangs mimic natural siRNAs and are commonly used but are not essential. Preferably, the overhangs should consist of two 5 nucleotides, most often dTdT or UU at the 3' end of the sense and antisense strand that are not complementary to the target sequence. The siRNAs may be produced by any method that is now known or that comes to be known for synthesizing double stranded RNA that one skilled in the art would appreciate would be useful in the present invention. Preferably, the siRNAs will be produced by Dharmacon's 10 proprietary ACE® technology. However, other methods for synthesizing siRNAs are well known to persons skilled in the art and include, but are not limited to, any chemical synthesis of RNA oligonucleotides, ligation of shorter oligonucleotides, *in vitro* transcription of RNA oligonucleotides, the use of vectors for expression within cells, recombinant Dicer products and PCR products.

15

The siRNA duplexes within the aforementioned pools of siRNAs may correspond to overlapping sequences within a particular mRNA, or non-overlapping sequences of the mRNA. However, preferably they correspond to non-overlapping sequences. Further, each siRNA may be selected randomly, or one or more of the 20 siRNA may be selected according to the criteria discussed above for maximizing the effectiveness of siRNA.

Included in the definition of siRNAs are siRNAs that contain substituted and/or labeled nucleotides that may, for example, be labeled by radioactivity, 25 fluorescence or mass. The most common substitutions are at the 2' position of the ribose sugar, where moieties such as H (hydrogen) F, NH<sub>3</sub>, OCH<sub>3</sub> and other O- alkyl, alkenyl, alkynyl, and orthoesters, may be substituted, or in the phosphorous backbone, where sulfur, amines or hydrocarbons may be substituted for the bridging of non-bridging atoms in the phosphodiester bond. Examples of modified siRNAs are 30 explained more fully in commonly assigned U.S. Patent Application Ser. No. 10/613,077, filed July 1, 2003, which is incorporated by reference herein.

Additionally, as noted above, the cell type into which the siRNA is introduced may affect the ability of the siRNA to enter the cell; however, it does not appear to

affect the ability of the siRNA to function once it enters the cell. Methods for introducing double-stranded RNA into various cell types are well known to persons skilled in the art.

5           As persons skilled in the art are aware, in certain species, the presence of proteins such as RdRP, the RNA-dependent RNA polymerase, may catalytically enhance the activity of the siRNA. For example, RdRP propagates the RNAi effect in *C. elegans* and other non-mammalian organisms. In fact, in organisms that contain these proteins, the siRNA may be inherited. Two other proteins that are well studied 10 and known to be a part of the machinery are members of the Argonaute family and Dicer, as well as their homologues. There is also initial evidence that the RISC complex might be associated with the ribosome so the more efficiently translated mRNAs will be more susceptible to silencing than others.

15           Another very important factor in the efficacy of siRNA is mRNA localization. In general, only cytoplasmic mRNAs are considered to be accessible to RNAi to any appreciable degree. However, appropriately designed siRNAs, for example, siRNAs modified with internucleotide linkages, may be able to cause silencing by acting in the nucleus. Examples of these types of modifications are described in commonly 20 assigned U.S. Patent Application Serial Nos. 10/431,027 and 10/613,077, each of which is incorporated by reference herein.

25           As described above, even when one selects at least two siRNAs at random, the effectiveness of the two may be greater than one would predict based on the effectiveness of two individual siRNAs. This additive or synergistic effect is particularly noticeable as one increases to at least three siRNAs, and even more noticeable as one moves to at least four siRNAs. Surprisingly, the pooling of the non-functional and semi-functional siRNAs, particularly more than five siRNAs, can lead to a silencing mixture that is as effective if not more effective than any one particular 30 functional siRNA.

Within the kit of the present invention, preferably each siRNA will be present in a concentration of between 0.001 and 200  $\mu$ M, more preferably between 0.01 and 200 nM, and most preferably between 0.1 and 10 nM.

5        In addition to preferably comprising at least four or five siRNAs, the kit of the present invention will also preferably comprise a buffer to keep the siRNA duplex stable. Persons skilled in the art are aware of buffers suitable for keeping siRNA stable. For example, the buffer may be comprised of 100 mM KCl, 30 mM HEPES-pH 7.5, and 1 mM MgCl<sub>2</sub>. Alternatively, kits might contain complementary strands 10 that contain any one of a number of chemical modifications (e.g. a 2'-O-ACE) that protect the agents from degradation by nucleases. In this instance, the user may (or may not) remove the modifying protective group (e.g. deprotect) before annealing the two complementary strands together.

15        By way of example, the kit may be organized such that pools of siRNA duplexes are provided on an array or microarray of wells or drops for a particular gene set or for unrelated genes. The array may, for example, be in 96 wells, 384 wells or 1284 wells arrayed in a plastic plate or on a glass slide using techniques now known or that come to be known to persons skilled in the art. Within an array, 20 preferably there will be controls such as functional anti-lamin A/C, cyclophilin and two siRNA duplexes that are not specific to the gene of interest.

25        In order to ensure stability of the siRNA pools prior to usage, they may be retained in lyophilized form at minus twenty degrees ( $-20^{\circ}\text{C}$ ) until they are ready for use. Prior to usage, they should be resuspended; however, even once resuspended, for example, in the aforementioned buffer, they should be kept at minus twenty degrees, ( $-20^{\circ}\text{C}$ ) until used. The aforementioned buffer, prior to use, may be stored at approximately 4°C or room temperature. Effective temperatures at which to conduct transfections are well known to persons skilled in the art and include for example, 30 room temperature.

The kit may be applied either *in vivo* or *in vitro*. Preferably, the siRNA of the pools or kits is applied to a cell through transfection, employing standard transfection

protocols. These methods are well known to persons skilled in the art and include the use of lipid-based carriers, electroporation, cationic carriers, and microinjection.

Further, one could apply the present invention by synthesizing equivalent DNA sequences (either as two separate, complementary strands, or as hairpin molecules)

5 instead of siRNA sequences and introducing them into cells through vectors. Once in the cells, the cloned DNA could be transcribed, thereby forcing the cells to generate the siRNA. Examples of vectors suitable for use with the present application include but are not limited to the standard transient expression vectors, adenoviruses, retroviruses, lentivirus-based vectors, as well as other traditional expression vectors.

10 Any vector that has an adequate siRNA expression and procession module may be used. Furthermore, certain chemical modifications to siRNAs, including but not limited to conjugations to other molecules, may be used to facilitate delivery. For certain applications it may be preferable to deliver molecules without transfection by simply formulating in a physiological acceptable solution.

15

This embodiment may be used in connection with any of the aforementioned embodiments. Accordingly, the sequences within any pool may be selected by rational design.

20

### **Multigene Silencing**

In addition to developing kits that contain multiple siRNA directed against a single gene, another embodiment includes the use of multiple siRNA targeting

multiple genes. Multiple genes may be targeted through the use of high- or hyper-

25 functional siRNA. High- or hyper- functional siRNA that exhibit increased potency, require lower concentrations to induce desired phenotypic (and thus therapeutic) effects. This circumvents RISC saturation. It therefore reasons that if lower concentrations of a single siRNA are needed for knockout or knockdown expression

of one gene, then the remaining (uncomplexed) RISC will be free and available to

30 interact with siRNA directed against two, three, four, or more, genes. Thus in this embodiment, the authors describe the use of highly functional or hyper-functional siRNA to knock out three separate genes. More preferably, such reagents could be combined to knockout four distinct genes. Even more preferably, highly functional or hyperfunctional siRNA could be used to knock out five distinct genes. Most

preferably, siRNA of this type could be used to knockout or knockdown the expression of six or more genes.

### Hyperfunctional siRNA

5        The term hyperfunctional siRNA (hf-siRNA) describes a subset of the siRNA population that induces RNAi in cells at low- or sub-nanomolar concentrations for extended periods of time. These traits, heightened potency and extended longevity of the RNAi phenotype, are highly attractive from a therapeutic standpoint. Agents having higher potency require lesser amounts of the molecule to achieve the desired 10 physiological response, thus reducing the probability of side effects due to "off-target" interference. In addition to the potential therapeutic benefits associated with hyperfunctional siRNA, hf-siRNA are also desirable from an economic position. Hyperfunctional siRNA may cost less on a per-treatment basis, thus reducing the overall expenditures to both the manufacturer and the consumer.

15        Identification of hyperfunctional siRNA involves multiple steps that are designed to examine an individual siRNA agent's concentration- and/or longevity-profiles. In one non-limiting example, a population of siRNA directed against a single gene are first analyzed using the previously described algorithm (Formula VIII). Individual siRNA are then introduced into a test cell line and assessed for the ability 20 to degrade the target mRNA. It is important to note that when performing this step it is not necessary to test all of the siRNA. Instead, it is sufficient to test only those siRNA having the highest SMARTscores<sup>TM</sup> (*i.e.* SMARTscore<sup>TM</sup> > -10). Subsequently, the gene silencing data is plotted against the SMARTscores<sup>TM</sup> (see 25 **Figure 9**). SiRNA that (1) induce a high degree of gene silencing (*i.e.* they induce greater than 80% gene knockdown) and (2) have superior SMARTscores<sup>TM</sup> (*i.e.* a SMARTscore<sup>TM</sup> of > -10, suggesting a desirable average internal stability profile) are selected for further investigations designed to better understand the molecule's potency and longevity. In one, non-limiting study dedicated to understanding a molecule's potency, an siRNA is introduced into one (or more) cell types in 30 increasingly diminishing concentrations (*e.g.* 3.0 → 0.3 nM). Subsequently, the level of gene silencing induced by each concentration is examined and siRNA that exhibit hyperfunctional potency (*i.e.* those that induce 80% silencing or greater at *e.g.* picomolar concentrations) are identified. In a second study, the longevity profiles of

siRNA having high (>-10) SMARTscores<sup>TM</sup> and greater than 80% silencing are examined. In one non-limiting example of how this is achieved, siRNA are introduced into a test cell line and the levels of RNAi are measured over an extended period of time (e.g. 24-168 hrs). SiRNAs that exhibit strong RNA interference patterns (*i.e.* 5 >80 % interference) for periods of time greater than, *e.g.*, 120 hours are thus identified. Studies similar to those described above can be performed on any and all of the >10<sup>6</sup> siRNA included in this document to further define the most functional molecule for any given gene. Molecules possessing one or both properties (extended longevity and heightened potency) are labeled “hyperfunctional siRNA,” and 10 earmarked as candidates for future therapeutic studies.

While the example(s) given above describe one means by which hyperfunctional siRNA can be isolated, neither the assays themselves nor the selection parameters used are rigid and can vary with each family of siRNA. Families 15 of siRNA include siRNAs directed against a single gene, or directed against a related family of genes.

The highest quality siRNA achievable for any given gene may vary considerably. Thus, for example, in the case of one gene (gene X), rigorous studies 20 such as those described above may enable the identification of an siRNA that, at picomolar concentrations, induces 99<sup>+</sup>% silencing for a period of 10 days. Yet identical studies of a second gene (gene Y) may yield an siRNA that at high nanomolar concentrations (*e.g.* 100nM) induces only 75% silencing for a period of 2 days. Both molecules represent the very optimum siRNA for their respective gene 25 targets and therefore are designated “hyperfunctional.” Yet due to a variety of factors including but not limited to target concentration, siRNA stability, cell type, off-target interference, and others, equivalent levels of potency and longevity are not achievable. Thus, for these reasons, the parameters described in the before mentioned assays, can vary. While the initial screen selected siRNA that had SMARTscores<sup>TM</sup> 30 above -10 and a gene silencing capability of greater than 80%, selections that have stronger (or weaker) parameters can be implemented. Similarly, in the subsequent studies designed to identify molecules with high potency and longevity, the desired cutoff criteria (*i.e.* the lowest concentration that induces a desirable level of interference, or the longest period of time that interference can be observed) can vary.

The experimentation subsequent to application of the rational criteria of this application is significantly reduced where one is trying to obtain a suitable hyperfunctional siRNA for, for example, therapeutic use. When, for example, the additional experimentation of the type described herein is applied by one skilled in the 5 art with this disclosure in hand, a hyperfunctional siRNA is readily identified.

The siRNA may be introduced into a cell by any method that is now known or that comes to be known and that from reading this disclosure, persons skilled in the art would determine would be useful in connection with the present invention in 10 enabling siRNA to cross the cellular membrane. These methods include, but are not limited to, any manner of transfection, such as for example transfection employing DEAE-Dextran, calcium phosphate, cationic lipids/liposomes, micelles, manipulation of pressure, microinjection, electroporation, immunoporation, use of vectors such as viruses, plasmids, cosmids, bacteriophages, cell fusions, and coupling of the 15 polynucleotides to specific conjugates or ligands such as antibodies, antigens, or receptors, passive introduction, adding moieties to the siRNA that facilitate its uptake, and the like.

Having described the invention with a degree of particularity, examples will 20 now be provided. These examples are not intended to and should not be construed to limit the scope of the claims in any way.

### Examples

#### **25 General Techniques and Nomenclatures**

siRNA nomenclature. All siRNA duplexes are referred to by sense strand. The first nucleotide of the 5'-end of the sense strand is position 1, which corresponds to position 19 of the antisense strand for a 19-mer. In most cases, to compare results from different experiments, silencing was determined by measuring specific transcript 30 mRNA levels or enzymatic activity associated with specific transcript levels, 24 hours post-transfection, with siRNA concentrations held constant at 100 nM. For all experiments, unless otherwise specified transfection efficiency was ensured to be over 95%, and no detectable cellular toxicity was observed. The following system of nomenclature was used to compare and report siRNA-silencing functionality: "F"

followed by the degree of minimal knockdown. For example, F50 signifies at least 50% knockdown, F80 means at least 80%, and so forth. For this study, all sub-F50 siRNAs were considered non-functional.

5    Cell culture and transfection. 96-well plates are coated with 50  $\mu$ l of 50 mg/ml poly-L-lysine (Sigma) for 1 hr, and then washed 3X with distilled water before being dried for 20 min. HEK293 cells or HEK293Lucs or any other cell type of interest are released from their solid support by trypsinization, diluted to  $3.5 \times 10^5$  cells/ml, followed by the addition of 100  $\mu$ L of cells/well. Plates are then incubated overnight 10 at 37° C, 5% CO<sub>2</sub>. Transfection procedures can vary widely depending on the cell type and transfection reagents. In one non-limiting example, a transfection mixture consisting of 2 mL Opti-MEM I (Gibco-BRL), 80  $\mu$ l Lipofectamine 2000 (Invitrogen), 15  $\mu$ L SUPERNasin at 20 U/ $\mu$ l (Ambion), and 1.5  $\mu$ l of reporter gene 15 plasmid at 1  $\mu$ g/  $\mu$ l is prepared in 5-ml polystyrene round bottom tubes. 100  $\mu$ l of transfection reagent is then combined with 100  $\mu$ l of siRNAs in polystyrene deep-well 20 titer plates (Beckman) and incubated for 20 to 30 min at room temp. 550  $\mu$ l of Opti-MEM is then added to each well to bring the final siRNA concentration to 100 nM. Plates are then sealed with parafilm and mixed. Media is removed from HEK293 cells and replaced with 95  $\mu$ l of transfection mixture. Cells are incubated overnight at 20 37° C, 5% CO<sub>2</sub>.

30    Quantification of gene knockdown. A variety of quantification procedures can be used to measure the level of silencing induced by siRNA or siRNA pools. In one non-limiting example: to measure mRNA levels 24 hrs post-transfection, QuantiGene branched-DNA (bDNA) kits (Bayer) (Wang, *et al*, *Regulation of insulin preRNA splicing by glucose*. Proc Natl Acad Sci 1997, 94:4360.) are used according to manufacturer instructions. To measure luciferase activity, media is removed from HEK293 cells 24 hrs post-transfection, and 50  $\mu$ l of Steady-GLO reagent (Promega) is added. After 5 min, plates are analyzed on a plate reader.

30

**Example I. Sequences Used to Develop the Algorithm.**

Anti-Firefly and anti-Cyclophilin siRNAs panels (**Figure 5a, b**) sorted according to using Formula VIII predicted values. All siRNAs scoring more than 0

(formula VIII) and more then 20 (formula IX) are fully functional. All ninety sequences for each gene (and DBI) appear below in **Table III**.

5

**TABLE III**

|       |    |              |                       |
|-------|----|--------------|-----------------------|
| Cyclo | 1  | SEQ. ID 0032 | GUUCCAAAAACAGUGGAUA   |
| Cyclo | 2  | SEQ. ID 0033 | UCCAAAAACAGUGGAUAAU   |
| Cyclo | 3  | SEQ. ID 0034 | CAAAAACAGUGGAUAAUUU   |
| Cyclo | 4  | SEQ. ID 0035 | AAAACAGUGGAUAAUUUUG   |
| Cyclo | 5  | SEQ. ID 0036 | AACAGUGGAUAAUUUUGUG   |
| Cyclo | 6  | SEQ. ID 0037 | CAGUGGAUAAUUUUGUGGC   |
| Cyclo | 7  | SEQ. ID 0038 | GUGGAUAAUUUUGUGGCCU   |
| Cyclo | 8  | SEQ. ID 0039 | GGAUAAUUUUGUGGCCUUA   |
| Cyclo | 9  | SEQ. ID 0040 | AUAAUUUUGUGGCCUUAGC   |
| Cyclo | 10 | SEQ. ID 0041 | AAUUUUGUGGCCUUAGCUA   |
| Cyclo | 11 | SEQ. ID 0042 | UUUUGUGGCCUUAGCUACA   |
| Cyclo | 12 | SEQ. ID 0043 | UUGUGGCCUUAGCUACAGG   |
| Cyclo | 13 | SEQ. ID 0044 | GUGGCCUUAGCUACAGGAG   |
| Cyclo | 14 | SEQ. ID 0045 | GGCCUUAGCUACAGGAGAG   |
| Cyclo | 15 | SEQ. ID 0046 | CCUUAGCUACAGGAGAGAA   |
| Cyclo | 16 | SEQ. ID 0047 | UUAGCUACAGGAGAGAAAG   |
| Cyclo | 17 | SEQ. ID 0048 | AGCUACAGGAGAGAAAGGA   |
| Cyclo | 18 | SEQ. ID 0049 | CUACAGGAGAGAAAGGAUU   |
| Cyclo | 19 | SEQ. ID 0050 | ACAGGAGAGAAAGGAUUUUG  |
| Cyclo | 20 | SEQ. ID 0051 | AGGAGAGAAAGGAUUUUGGC  |
| Cyclo | 21 | SEQ. ID 0052 | GAGAGAAAGGAUUUUGGCUA  |
| Cyclo | 22 | SEQ. ID 0053 | GAGAAAGGAUUUUGGCUACA  |
| Cyclo | 23 | SEQ. ID 0054 | GAAAGGAUUUUGGCUACAAA  |
| Cyclo | 24 | SEQ. ID 0055 | AAGGAUUUUGGCUACAAAAAA |
| Cyclo | 25 | SEQ. ID 0056 | GGAUUUGGCUACAAAAACAA  |
| Cyclo | 26 | SEQ. ID 0057 | AUUUGGCUACAAAAACAGC   |
| Cyclo | 27 | SEQ. ID 0058 | UUGGCUACAAAAACAGCAA   |
| Cyclo | 28 | SEQ. ID 0059 | GGCUACAAAAACAGCAAUAU  |
| Cyclo | 29 | SEQ. ID 0060 | CUACAAAAACAGCAAUUC    |
| Cyclo | 30 | SEQ. ID 0061 | ACAAAAACAGCAAUUCCA    |
| Cyclo | 31 | SEQ. ID 0062 | AAAAACAGCAAUUCCAUC    |
| Cyclo | 32 | SEQ. ID 0063 | AAACAGCAAUUCCAUCGU    |
| Cyclo | 33 | SEQ. ID 0064 | ACAGCAAUUCCAUCGUGU    |
| Cyclo | 34 | SEQ. ID 0065 | AGCAAUUCCAUCGUGUAA    |

|       |    |              |                      |
|-------|----|--------------|----------------------|
| Cyclo | 35 | SEQ. ID 0066 | CAAAUCCAUCGUGUAAUC   |
| Cyclo | 36 | SEQ. ID 0067 | AAUCCAUCGUGUAAUCAA   |
| Cyclo | 37 | SEQ. ID 0068 | UUCCAUCGUGUAAUCAAGG  |
| Cyclo | 38 | SEQ. ID 0069 | CCAUCGUGUAAUCAAGGAC  |
| Cyclo | 39 | SEQ. ID 0070 | AUCGUGUAAUCAAGGACUU  |
| Cyclo | 40 | SEQ. ID 0071 | CGUGUAAUCAAGGACUUCA  |
| Cyclo | 41 | SEQ. ID 0072 | UGUAAUCAAGGACUUCAUG  |
| Cyclo | 42 | SEQ. ID 0073 | UAAUCAAGGACUUCAUGAU  |
| Cyclo | 43 | SEQ. ID 0074 | AUCAAGGACUUCAUGAUCC  |
| Cyclo | 44 | SEQ. ID 0075 | CAAGGACUUCAUGAUCCAG  |
| Cyclo | 45 | SEQ. ID 0076 | AGGACUUCAUGAUCCAGGG  |
| Cyclo | 46 | SEQ. ID 0077 | GACUUCAUGAUCCAGGGCG  |
| Cyclo | 47 | SEQ. ID 0078 | CUUCAUGAUCCAGGGCGGA  |
| Cyclo | 48 | SEQ. ID 0079 | UCAUGAUCCAGGGCGGAGA  |
| Cyclo | 49 | SEQ. ID 0080 | AUGAUCCAGGGCGGAGACU  |
| Cyclo | 50 | SEQ. ID 0081 | GAUCCAGGGCGGAGACUUC  |
| Cyclo | 51 | SEQ. ID 0082 | UCCAGGGCGGAGACUUCAC  |
| Cyclo | 52 | SEQ. ID 0083 | CAGGGCGGAGACUUCACCA  |
| Cyclo | 53 | SEQ. ID 0084 | GGGCGGAGACUUCACCAGG  |
| Cyclo | 54 | SEQ. ID 0085 | GC GGAGACUUCACCAGGGG |
| Cyclo | 55 | SEQ. ID 0086 | GGAGACUUCACCAGGGGAG  |
| Cyclo | 56 | SEQ. ID 0087 | AGACUUCACCAGGGGAGAU  |
| Cyclo | 57 | SEQ. ID 0088 | ACUUCACCAGGGGAGAUGG  |
| Cyclo | 58 | SEQ. ID 0089 | UUCACCAGGGGAGAUGGCA  |
| Cyclo | 59 | SEQ. ID 0090 | CACCAGGGGAGAUGGCACA  |
| Cyclo | 60 | SEQ. ID 0091 | CCAGGGGAGAUGGCACAGG  |
| Cyclo | 61 | SEQ. ID 0092 | AGGGGAGAUGGCACAGGAG  |
| Cyclo | 62 | SEQ. ID 0093 | GGGAGAUGGCACAGGAGGA  |
| Cyclo | 63 | SEQ. ID 0094 | GAGAUGGCACAGGAGGAAA  |
| Cyclo | 64 | SEQ. ID 0095 | GAUGGCACAGGAGGAAAGA  |
| Cyclo | 65 | SEQ. ID 0094 | UGGCACAGGAGGAAAGAGC  |
| Cyclo | 66 | SEQ. ID 0096 | GCACAGGAGGAAAGAGCAU  |
| Cyclo | 67 | SEQ. ID 0097 | ACAGGAGGAAAGAGCAUCU  |
| Cyclo | 68 | SEQ. ID 0098 | AGGAGGAAAGAGCAUCUAC  |
| Cyclo | 69 | SEQ. ID 0099 | GAGGAAAGAGCAUCUACGG  |
| Cyclo | 70 | SEQ. ID 0100 | GGAAAGAGCAUCUACGGUG  |
| Cyclo | 71 | SEQ. ID 0101 | AAAGAGCAUCUACGGUGAG  |
| Cyclo | 72 | SEQ. ID 0102 | AGAGCAUCUACGGUGAGCG  |
| Cyclo | 73 | SEQ. ID 0103 | AGCAUCUACGGUGAGCGCU  |

|       |    |              |                      |
|-------|----|--------------|----------------------|
| Cyclo | 74 | SEQ. ID 0104 | CAUCUACGGUGAGCGCUUC  |
| Cyclo | 75 | SEQ. ID 0105 | UCUACGGUGAGCGCUUCCCC |
| Cyclo | 76 | SEQ. ID 0106 | UACGGUGAGCGCUUCCCCG  |
| Cyclo | 77 | SEQ. ID 0107 | CGGUGAGCGCUUCCCCGAU  |
| Cyclo | 78 | SEQ. ID 0108 | GUGAGCGCUUCCCCGAUGA  |
| Cyclo | 79 | SEQ. ID 0109 | GAGCGCUUCCCCGAUGAGA  |
| Cyclo | 80 | SEQ. ID 0110 | GCGCUUCCCCGAUGAGAAC  |
| Cyclo | 81 | SEQ. ID 0111 | GCUUCCCCGAUGAGAACUU  |
| Cyclo | 82 | SEQ. ID 0112 | UUCCCCGAUGAGAACUUCA  |
| Cyclo | 83 | SEQ. ID 0113 | CCCCGAUGAGAACUUCAAA  |
| Cyclo | 84 | SEQ. ID 0114 | CCGAUGAGAACUUCAAACU  |
| Cyclo | 85 | SEQ. ID 0115 | GAUGAGAACUUCAAACUGA  |
| Cyclo | 86 | SEQ. ID 0116 | UGAGAACUUCAAACUGAAG  |
| Cyclo | 87 | SEQ. ID 0117 | AGAACUUCAAACUGAAGCA  |
| Cyclo | 88 | SEQ. ID 0118 | AACUUCAAACUGAAGCACU  |
| Cyclo | 89 | SEQ. ID 0119 | CUUCAAACUGAAGCACUAC  |
| Cyclo | 90 | SEQ. ID 0120 | UCAAACUGAAGCACUACGG  |
| DB    | 1  | SEQ. ID 0121 | ACGGGCAAGGCCAAGUGGG  |
| DB    | 2  | SEQ. ID 0122 | CGGGCAAGGCCAAGUGGGGA |
| DB    | 3  | SEQ. ID 0123 | GGGCAAGGCCAAGUGGGGAU |
| DB    | 4  | SEQ. ID 0124 | GGCAAGGCCAAGUGGGGAUG |
| DB    | 5  | SEQ. ID 0125 | GCAAGGCCAAGUGGGGAUGC |
| DB    | 6  | SEQ. ID 0126 | CAAGGCCAAGUGGGGAUGCC |
| DB    | 7  | SEQ. ID 0127 | AAGGCCAAGUGGGGAUGCCU |
| DB    | 8  | SEQ. ID 0128 | AGGCCAAGUGGGGAUGCCUG |
| DB    | 9  | SEQ. ID 0129 | GGCCAAGUGGGGAUGCCUGG |
| DB    | 10 | SEQ. ID 0130 | GCCAAGUGGGGAUGCCUGGA |
| DB    | 11 | SEQ. ID 0131 | CCAAGUGGGGAUGCCUGGAA |
| DB    | 12 | SEQ. ID 0132 | CAAGUGGGGAUGCCUGGAAU |
| DB    | 13 | SEQ. ID 0133 | AAGUGGGGAUGCCUGGAAUG |
| DB    | 14 | SEQ. ID 0134 | AGUGGGGAUGCCUGGAAUGA |
| DB    | 15 | SEQ. ID 0135 | GUGGGGAUGCCUGGAAUGAG |
| DB    | 16 | SEQ. ID 0136 | UGGGGAUGCCUGGAAUGAGC |
| DB    | 17 | SEQ. ID 0137 | GGGAUGCCUGGAAUGAGCU  |
| DB    | 18 | SEQ. ID 0138 | GGAUGCCUGGAAUGAGCUG  |
| DB    | 19 | SEQ. ID 0139 | GAUGCCUGGAAUGAGCUGA  |
| DB    | 20 | SEQ. ID 0140 | AUGCCUGGAAUGAGCUGAA  |
| DB    | 21 | SEQ. ID 0141 | UGCCUGGAAUGAGCUGAAA  |
| DB    | 22 | SEQ. ID 0142 | GCCUGGAAUGAGCUGAAAG  |
| DB    | 23 | SEQ. ID 0143 | CCUGGAAUGAGCUGAAAGG  |

|    |    |              |                      |
|----|----|--------------|----------------------|
| DB | 24 | SEQ. ID 0144 | CUGGAAUGAGCUGAAAGGG  |
| DB | 25 | SEQ. ID 0145 | UGGAAUGAGCUGAAAGGGAA |
| DB | 26 | SEQ. ID 0146 | GGAAUGAGCUGAAAGGGAC  |
| DB | 27 | SEQ. ID 0147 | GAAUGAGCUGAAAGGGACU  |
| DB | 28 | SEQ. ID 0148 | AAUGAGCUGAAAGGGACUU  |
| DB | 29 | SEQ. ID 0149 | AUGAGCUGAAAGGGACUUC  |
| DB | 30 | SEQ. ID 0150 | UGAGCUGAAAGGGACUUCC  |
| DB | 31 | SEQ. ID 0151 | GAGCUGAAAGGGACUUC    |
| DB | 32 | SEQ. ID 0152 | AGCUGAAAGGGACUUC     |
| DB | 33 | SEQ. ID 0153 | GCUGAAAGGGACUUC      |
| DB | 34 | SEQ. ID 0154 | CUGAAAGGGACUUC       |
| DB | 35 | SEQ. ID 0155 | UGAAAGGGACUUC        |
| DB | 36 | SEQ. ID 0156 | GAAAGGGACUUC         |
| DB | 37 | SEQ. ID 0157 | AAAGGGACUUC          |
| DB | 38 | SEQ. ID 0158 | AAGGGACUUC           |
| DB | 39 | SEQ. ID 0159 | AGGGACUUC            |
| DB | 40 | SEQ. ID 0160 | GGGACUUC             |
| DB | 41 | SEQ. ID 0161 | GGACUUC              |
| DB | 42 | SEQ. ID 0162 | GACUUC               |
| DB | 43 | SEQ. ID 0163 | ACUUCC               |
| DB | 44 | SEQ. ID 0164 | CUUCC                |
| DB | 45 | SEQ. ID 0165 | UCCCAAGGAAAGAUGCCAUG |
| DB | 46 | SEQ. ID 0166 | UCCAAGGAAAGAUGCCAUGA |
| DB | 47 | SEQ. ID 0167 | CCAAGGAAAGAUGCCAUGAA |
| DB | 48 | SEQ. ID 0168 | CAAGGAAAGAUGCCAUGAAA |
| DB | 49 | SEQ. ID 0169 | AAGGAAAGAUGCCAUGAAAG |
| DB | 50 | SEQ. ID 0170 | AGGAAGAUGCCAUGAAAGC  |
| DB | 51 | SEQ. ID 0171 | GGAAGAUGCCAUGAAAGCU  |
| DB | 52 | SEQ. ID 0172 | GAAGAUGCCAUGAAAGCUU  |
| DB | 53 | SEQ. ID 0173 | AAGAUGCCAUGAAAGCUUA  |
| DB | 54 | SEQ. ID 0174 | AGAUGCCAUGAAAGCUUAC  |
| DB | 55 | SEQ. ID 0175 | GAUGCCAUGAAAGCUUACA  |
| DB | 56 | SEQ. ID 0176 | AUGCCAUGAAAGCUUACAU  |
| DB | 57 | SEQ. ID 0177 | UGCCAUGAAAGCUUACAUCA |
| DB | 58 | SEQ. ID 0178 | GCCAUGAAAGCUUACAUCA  |
| DB | 59 | SEQ. ID 0179 | CCAUGAAAGCUUACAUCAA  |
| DB | 60 | SEQ. ID 0180 | CAUGAAAGCUUACAUCAAC  |
| DB | 61 | SEQ. ID 0181 | AUGAAAGCUUACAUCAACA  |
| DB | 62 | SEQ. ID 0182 | UGAAAGCUUACAUCAACAA  |
| DB | 63 | SEQ. ID 0183 | GAAAGCUUACAUCAACAA   |

|     |    |              |                      |
|-----|----|--------------|----------------------|
| DB  | 64 | SEQ. ID 0184 | AAAGCUUACAUCAACAAAG  |
| DB  | 65 | SEQ. ID 0185 | AAGCUUACAUCAACAAAGU  |
| DB  | 66 | SEQ. ID 0186 | AGCUUACAUCAACAAAGUA  |
| DB  | 67 | SEQ. ID 0187 | GCUUACAUCAACAAAGUAG  |
| DB  | 68 | SEQ. ID 0188 | CUUACAUCAACAAAGUAGA  |
| DB  | 69 | SEQ. ID 0189 | UUACAUCAACAAAGUAGAA  |
| DB  | 70 | SEQ. ID 0190 | UACAUCAACAAAGUAGAAG  |
| DB  | 71 | SEQ. ID 0191 | ACAUCAACAAAGUAGAAGA  |
| DB  | 72 | SEQ. ID 0192 | CAUCAACAAAGUAGAAGAG  |
| DB  | 73 | SEQ. ID 0193 | AUCAACAAAGUAGAAGAGC  |
| DB  | 74 | SEQ. ID 0194 | UCAACAAAGUAGAAGAGCU  |
| DB  | 75 | SEQ. ID 0195 | CAACAAAGUAGAAGAGCUA  |
| DB  | 76 | SEQ. ID 0196 | AACAAAGUAGAAGAGCUAA  |
| DB  | 77 | SEQ. ID 0197 | ACAAAGUAGAAGAGCUAAA  |
| DB  | 78 | SEQ. ID 0198 | CAAAGUAGAAGAGCUAAAG  |
| DB  | 79 | SEQ. ID 0199 | AAAGUAGAAGAGCUAAAGA  |
| DB  | 80 | SEQ. ID 0200 | AAGUAGAAGAGCUAAAGAA  |
| DB  | 81 | SEQ. ID 0201 | AGUAGAAGAGCUAAAGAAA  |
| DB  | 82 | SEQ. ID 0202 | GUAGAAGAGCUAAAGAAAA  |
| DB  | 83 | SEQ. ID 0203 | UAGAAGAGCUAAAGAAAAAA |
| DB  | 84 | SEQ. ID 0204 | AGAAGAGCUAAAGAAAAAAA |
| DB  | 85 | SEQ. ID 0205 | GAAGAGCUAAAGAAAAAAAU |
| DB  | 86 | SEQ. ID 0206 | AAGAGCUAAAGAAAAAAUA  |
| DB  | 87 | SEQ. ID 0207 | AGAGCUAAAGAAAAAAUAC  |
| DB  | 88 | SEQ. ID 0208 | GAGCUAAAGAAAAAAUACG  |
| DB  | 89 | SEQ. ID 0209 | AGCUAAAGAAAAAAUACGG  |
| DB  | 90 | SEQ. ID 0210 | GCUAAGAAAAAAUACGGG   |
| Luc | 1  | SEQ. ID 0211 | AUCCUCAUAAAGGCCAAGA  |
| Luc | 2  | SEQ. ID 0212 | AGAUCCUCAUAAAGGCCAA  |
| Luc | 3  | SEQ. ID 0213 | AGAGAUCCUCAUAAAGGCC  |
| Luc | 4  | SEQ. ID 0214 | AGAGAGAUCCUCAUAAAGG  |
| Luc | 5  | SEQ. ID 0215 | UCAGAGAGAUCCUCAUAAA  |
| Luc | 6  | SEQ. ID 0216 | AAUCAGAGAGAUCCUCUA   |
| Luc | 7  | SEQ. ID 0217 | AAAAUCAGAGAGAUCCUCA  |
| Luc | 8  | SEQ. ID 0218 | GAAAAAUCAGAGAGAUCCU  |
| Luc | 9  | SEQ. ID 0219 | AAGAAAAAUCAGAGAGAUC  |
| Luc | 10 | SEQ. ID 0220 | GCAAGAAAAAUCAGAGAGA  |
| Luc | 11 | SEQ. ID 0221 | ACGCAAGAAAAAUCAGAGA  |
| Luc | 12 | SEQ. ID 0222 | CGACGCAAGAAAAAUCAGA  |
| Luc | 13 | SEQ. ID 0223 | CUCGACGCAAGAAAAAUCA  |

|     |    |              |                      |
|-----|----|--------------|----------------------|
| Luc | 14 | SEQ. ID 0224 | AACUCGACGCAAGAAAAAU  |
| Luc | 15 | SEQ. ID 0225 | AAAACUCGACGCAAGAAAA  |
| Luc | 16 | SEQ. ID 0226 | GGAAAACUCGACGCAAGAA  |
| Luc | 17 | SEQ. ID 0227 | CCGGAAAACUCGACGCAAG  |
| Luc | 18 | SEQ. ID 0228 | UACCGGAAAACUCGACGCA  |
| Luc | 19 | SEQ. ID 0229 | CUUACCGGAAAACUCGACG  |
| Luc | 20 | SEQ. ID 0230 | GUCUUACCGGAAAACUCGA  |
| Luc | 21 | SEQ. ID 0231 | AGGUCUUACCGGAAAACUC  |
| Luc | 22 | SEQ. ID 0232 | AAAGGUCUUACCGGAAAAC  |
| Luc | 23 | SEQ. ID 0233 | CGAAAGGUCUUACCGGAAA  |
| Luc | 24 | SEQ. ID 0234 | ACCGAAAGGUCUUACCGGA  |
| Luc | 25 | SEQ. ID 0235 | GUACCGAAAGGUCUUACCG  |
| Luc | 26 | SEQ. ID 0236 | AAGUACCGAAAGGUCUUAC  |
| Luc | 27 | SEQ. ID 0237 | CGAAGUACCGAAAGGUCUU  |
| Luc | 28 | SEQ. ID 0238 | GACGAAGUACCGAAAGGUC  |
| Luc | 29 | SEQ. ID 0239 | UGGACGAAGUACCGAAAGG  |
| Luc | 30 | SEQ. ID 0240 | UGUGGACGAAGUACCGAAA  |
| Luc | 31 | SEQ. ID 0241 | UUUGUGGACGAAGUACCGA  |
| Luc | 32 | SEQ. ID 0242 | UGUUUUGUGGACGAAGUACC |
| Luc | 33 | SEQ. ID 0243 | UGUGUUUGUGGACGAAGUA  |
| Luc | 34 | SEQ. ID 0244 | GUUGUGUUUGUGGACGAAG  |
| Luc | 35 | SEQ. ID 0245 | GAGUUGUGUUUGUGGACGA  |
| Luc | 36 | SEQ. ID 0246 | AGGAGUUGUGUUUGUGGAC  |
| Luc | 37 | SEQ. ID 0247 | GGAGGAGUUGUGUUUGUGG  |
| Luc | 38 | SEQ. ID 0248 | GCGGAGGAGUUGUGUUUGU  |
| Luc | 39 | SEQ. ID 0249 | GCGCAGGAGGAGUUGUGUUU |
| Luc | 40 | SEQ. ID 0250 | UUGCGCGAGGAGUUGUGU   |
| Luc | 41 | SEQ. ID 0251 | AGUUGCGCGAGGAGUUGUGU |
| Luc | 42 | SEQ. ID 0252 | AAAGUUGCGCGAGGAGUU   |
| Luc | 43 | SEQ. ID 0253 | AAAAAGUUGCGCGAGGAGAG |
| Luc | 44 | SEQ. ID 0254 | CGAAAAAGUUGCGCGAGGAG |
| Luc | 45 | SEQ. ID 0255 | CGCGAAAAAGUUGCGCGAGG |
| Luc | 46 | SEQ. ID 0256 | ACCGCGAAAAAGUUGCGCG  |
| Luc | 47 | SEQ. ID 0257 | CAACCGCGAAAAAGUUGCG  |
| Luc | 48 | SEQ. ID 0258 | AACAACCGCGAAAAAGUUG  |
| Luc | 49 | SEQ. ID 0259 | GUAACAACCGCGAAAAAGU  |
| Luc | 50 | SEQ. ID 0260 | AAGUAACAACCGCGAAAAAA |
| Luc | 51 | SEQ. ID 0261 | UCAAGUAACAACCGCGAAA  |
| Luc | 52 | SEQ. ID 0262 | AGUCAAGUAACAACCGCGA  |
| Luc | 53 | SEQ. ID 0263 | CCAGUCAAGUAACAACCGC  |

|     |    |              |                     |
|-----|----|--------------|---------------------|
| Luc | 54 | SEQ. ID 0264 | CGCCAGUCAAGUAACAACC |
| Luc | 55 | SEQ. ID 0265 | GUCGCCAGUCAAGUAACAA |
| Luc | 56 | SEQ. ID 0266 | ACGUCGCCAGUCAAGUAAC |
| Luc | 57 | SEQ. ID 0267 | UUACGUCGCCAGUCAAGUA |
| Luc | 58 | SEQ. ID 0268 | GAUUACGUCGCCAGUCAAG |
| Luc | 59 | SEQ. ID 0269 | UGGAUUACGUCGCCAGUCA |
| Luc | 60 | SEQ. ID 0270 | CGUGGAUUACGUCGCCAGU |
| Luc | 61 | SEQ. ID 0271 | AUCGUGGAUUACGUCGCCA |
| Luc | 62 | SEQ. ID 0272 | AGAUCGUGGAUUACGUCGC |
| Luc | 63 | SEQ. ID 0273 | AGAGAUCGUGGAUUACGUC |
| Luc | 64 | SEQ. ID 0274 | AAAGAGAUCGUGGAUUACG |
| Luc | 65 | SEQ. ID 0275 | AAAAAGAGAUCGUGGAUUA |
| Luc | 66 | SEQ. ID 0276 | GGAAAAAGAGAUCGUGGAU |
| Luc | 67 | SEQ. ID 0277 | ACGGAAAAAGAGAUCGUGG |
| Luc | 68 | SEQ. ID 0278 | UGACGGAAAAAGAGAUCGU |
| Luc | 69 | SEQ. ID 0279 | GAUGACGGAAAAAGAGAUC |
| Luc | 70 | SEQ. ID 0280 | ACGAUGACGGAAAAAGAGA |
| Luc | 71 | SEQ. ID 0281 | AGACGAUGACGGAAAAAGA |
| Luc | 72 | SEQ. ID 0282 | AAAGACGAUGACGGAAAAA |
| Luc | 73 | SEQ. ID 0283 | GGAAAGACGAUGACGGAAA |
| Luc | 74 | SEQ. ID 0284 | ACGGAAAGACGAUGACGGA |
| Luc | 75 | SEQ. ID 0285 | GCACGGAAAGACGAUGACG |
| Luc | 76 | SEQ. ID 0286 | GAGCACGGAAAGACGAUGA |
| Luc | 77 | SEQ. ID 0287 | UGGAGCACGGAAAGACGAU |
| Luc | 78 | SEQ. ID 0288 | UUUGGAGCACGGAAAGACG |
| Luc | 79 | SEQ. ID 0289 | GUUUUGGAGCACGGAAAGA |
| Luc | 80 | SEQ. ID 0290 | UUGUUUUGGAGCACGGAAA |
| Luc | 81 | SEQ. ID 0291 | UGUUGUUUUGGAGCACGGA |
| Luc | 82 | SEQ. ID 0292 | GUUGUUGUUUUGGAGCACG |
| Luc | 83 | SEQ. ID 0293 | CCGUUGUUGUUUUGGAGCA |
| Luc | 84 | SEQ. ID 0294 | CGCCGUUGUUGUUUUGGAG |
| Luc | 85 | SEQ. ID 0295 | GCCGCCGUUGUUGUUUUGG |
| Luc | 86 | SEQ. ID 0296 | CCGCCGCCGUUGUUGUUUU |
| Luc | 87 | SEQ. ID 0297 | UCCCGCCGCCGUUGUUGUU |
| Luc | 88 | SEQ. ID 0298 | CUUCCCGCCGCCGUUGUUG |
| Luc | 89 | SEQ. ID 0299 | AACUUCCCGCCGCCGUUGU |
| Luc | 90 | SEQ. ID 0300 | UGAACUUCCCGCCGCCGUU |

**Example II. Validation of the Algorithm using DBI, Luciferase, PLK, EGFR, and SEAP**

The algorithm (Formula VIII) identified siRNAs for five genes, human DBI, firefly luciferase (fLuc), renilla luciferase (rLuc), human PLK, and human secreted 5 alkaline phosphatase (SEAP). Four individual siRNAs were selected on the basis of their SMARTscores™ derived by analysis of their sequence using Formula VIII (all of the siRNAs would be selected with Formula IX as well) and analyzed for their ability to silence their targets' expression. In addition to the scoring, a BLAST search was conducted for each siRNA. To minimize the potential for off-target silencing 10 effects, only those target sequences with more than three mismatches against unrelated sequences were selected. Semizarov, *et al*, *Specificity of short interfering RNA determined through gene expression signatures*. Proc. Natl. Acad. Sci. U.S.A. 2003, 100:6347. These duplexes were analyzed individually and in pools of 4 and compared with several siRNAs that were randomly selected. The functionality was 15 measured a percentage of targeted gene knockdown as compared to controls. All siRNAs were transfected as described by the methods above at 100 nM concentration into HEK293 using Lipofectamine 2000. The level of the targeted gene expression was evaluated by B-DNA as described above and normalized to the non-specific control. **Figure 10** shows that the siRNAs selected by the algorithm disclosed herein 20 were significantly more potent than randomly selected siRNAs. The algorithm increased the chances of identifying an F50 siRNA from 48% to 91%, and an F80 siRNA from 13% to 57%. In addition, pools of SMART siRNA silence the selected target better than randomly selected pools (see Figure 10F).

25 **Example III. Validation of the Algorithm Using Genes Involved in Clathrin-Dependent Endocytosis.**

Components of clathrin-mediated endocytosis pathway are key to modulating intracellular signaling and play important roles in disease. Chromosomal rearrangements that result in fusion transcripts between the Mixed-Lineage Leukemia 30 gene (MLL) and CALM (Clathrin assembly lymphoid myeloid leukemia gene) are believed to play a role in leukemogenesis. Similarly, disruptions in Rab7 and Rab9, as well as HIP1 (Huntingtin-interacting protein), genes that are believed to be involved in endocytosis, are potentially responsible for ailments resulting in lipid storage, and neuronal diseases, respectively. For these reasons, siRNA directed

against clathrin and other genes involved in the clathrin-mediated endocytotic pathway are potentially important research and therapeutic tools.

5 siRNAs directed against genes involved in the clathrin-mediated endocytosis pathways were selected using Formula VIII. The targeted genes were clathrin heavy chain (CHC, accession # NM\_004859), clathrin light chain A (CLCa, NM\_001833), clathrin light chain B (CLCb, NM\_001834), CALM (U45976),  $\beta$ 2 subunit of AP-2 ( $\beta$ 2, NM\_001282), Eps15 (NM\_001981), Eps15R (NM\_021235), dynamin II (DYNII, NM\_004945), Rab5a (BC001267), Rab5b (NM\_002868), Rab5c (AF141304), and EEA.1 (XM\_018197).

10 For each gene, four siRNAs duplexes with the highest scores were selected and a BLAST search was conducted for each of them using the Human EST database. In order to minimize the potential for off-target silencing effects, only those 15 sequences with more than three mismatches against un-related sequences were used. All duplexes were synthesized at Dharmacon, Inc. as 21-mers with 3'-UU overhangs using a modified method of 2'-ACE chemistry Scaringe, *Advanced 5'-silyl-2'-orthoester approach to RNA oligonucleotide synthesis*, Methods Enzymol 2000, 317:3 and the antisense strand was chemically phosphorylated to insure maximized activity.

20

HeLa cells were grown in Dulbecco's modified Eagle's medium (DMEM) containing 10% fetal bovine serum, antibiotics and glutamine. siRNA duplexes were resuspended in 1X siRNA Universal buffer (Dharmacon, Inc.) to 20 $\mu$ M prior to transfection. HeLa cells in 12-well plates were transfected twice with 4 $\mu$ l of 20 $\mu$ M 25 siRNA duplex in 3 $\mu$ l Lipofectamine 2000 reagent (Invitrogen, Carlsbad, California, USA) at 24-hour intervals. For the transfections in which 2 or 3 siRNA duplexes were included, the amount of each duplex was decreased, so that the total amount was the same as in transfections with single siRNAs. Cells were plated into normal culture medium 12 hours prior to experiments, and protein levels were measured 2 or 4 days 30 after the first transfection.

Equal amounts of lysates were resolved by electrophoresis, blotted, and stained with the antibody specific to targeted protein, as well as antibodies specific to

unrelated proteins, PP1 phosphatase and Tsg101 (not shown). The cells were lysed in Triton X-100/glycerol solubilization buffer as described previously. Tebar, Bohlander, & Sorkin, *Clathrin Assembly Lymphoid Myeloid Leukemia (CALM) Protein: Localization in Endocytic-coated Pits, Interactions with Clathrin, and the Impact of Overexpression on Clathrin-mediated Traffic*, Mol. Biol. Cell Aug 1999, 10:2687. Cell lysates were electrophoresed, transferred to nitrocellulose membranes, and Western blotting was performed with several antibodies followed by detection using enhanced chemiluminescence system (Pierce, Inc). Several x-ray films were analyzed to determine the linear range of the chemiluminescence signals, and the quantifications were performed using densitometry and AlphaImager v5.5 software (Alpha Innotech Corporation). In experiments with Eps15R-targeted siRNAs, cell lysates were subjected to immunoprecipitation with Ab860, and Eps15R was detected in immunoprecipitates by Western blotting as described above.

15 The antibodies to assess the levels of each protein by Western blot were obtained from the following sources: monoclonal antibody to clathrin heavy chain (TD.1) was obtained from American Type Culture Collection (Rockville, MD, USA); polyclonal antibody to dynamin II was obtained from Affinity Bioreagents, Inc. (Golden, CO, USA); monoclonal antibodies to EEA.1 and Rab5a were purchased from BD Transduction Laboratories (Los Angeles, CA, USA); the monoclonal antibody to Tsg101 was purchased from Santa Cruz Biotechnology, Inc. (Santa Cruz, CA, USA); the monoclonal antibody to GFP was from ZYMED Laboratories Inc. (South San Francisco, CA, USA); the rabbit polyclonal antibodies Ab32 specific to  $\alpha$ -adaptins and Ab20 to CALM were described previously Sorkin, *et al, Stoichiometric Interaction of the Epidermal Growth Factor Receptor with the Clathrin-associated Protein Complex AP-2*, J. Biol. Chem. Jan 1995, 270:619, the polyclonal antibodies to clathrin light chains A and B were kindly provided by Dr. F. Brodsky (UCSF); monoclonal antibodies to PP1 (BD Transduction Laboratories) and  $\alpha$ -Actinin (Chemicon) were kindly provided by Dr. M. Dell'Acqua (University of Colorado); 20 Eps15 Ab577 and Eps15R Ab860 were kindly provided by Dr. P.P. Di Fiore (European Cancer Institute).

25

Figure 11 demonstrates the *in vivo* functionality of 48 individual siRNAs, selected using Formula VIII (most of them will meet the criteria incorporated by Formula IX as well) targeting 12 genes. Various cell lines were transfected with siRNA duplexes (*Dup1-4*) or pools of siRNA duplexes (Pool), and the cells were lysed 3 days after transfection with the exception of CALM (2 days) and  $\beta$ 2 (4 days).

Note a  $\beta$ 1-adaptin band (part of AP-1 Golgi adaptor complex) that runs slightly slower than  $\beta$ 2 adaptin. CALM has two splice variants, 66 and 72 kD. The full-length Eps15R (a doublet of ~130 kD) and several truncated spliced forms of ~ 100 kD and ~70 kD were detected in Eps15R immunoprecipitates (shown by arrows). The cells were lysed 3 days after transfection. Equal amounts of lysates were resolved by electrophoresis and blotted with the antibody specific to a targeted protein (GFP antibody for YFP fusion proteins) and the antibody specific to unrelated proteins PP1 phosphatase or  $\alpha$ -actinin, and TSG101. The amount of protein in each specific band was normalized to the amount of non-specific proteins in each lane of the gel. Nearly all of them appear to be functional, which establishes that Formula VIII and IX can be used to predict siRNAs' functionality in general in a genome wide manner.

To generate the fusion of yellow fluorescent protein (YFP) with Rab5b or Rab5c (YFP-Rab5b or YFP-Rab5c), a DNA fragment encoding the full-length human Rab5b or Rab5c was obtained by PCR using *Pfu* polymerase (Stratagene) with a *Sac*I restriction site introduced into the 5' end and a *Kpn*I site into the 3' end and cloned into pEYFP-C1 vector (CLONTECH, Palo Alto, CA, USA). GFP-CALM and YFP-Rab5a were described previously Tebar, Bohlander, & Sorkin, *Clathrin Assembly Lymphoid Myeloid Leukemia (CALM) Protein: Localization in Endocytic-coated Pits, Interactions with Clathrin, and the Impact of Overexpression on Clathrin-mediated Traffic*, Mol. Biol. Cell Aug 1999, 10:2687.

Example III. Validation of the Algorithm Using Eg5, GADPH, ATE1, MEK2, MEK1, QB, LaminA/C, c-myc, human cyclophilin, and mouse cyclophilin.

A number of genes have been identified as playing potentially important roles in disease etiology. Expression profiles of normal and diseased kidneys has implicated Edg5 in immunoglobulin A neuropathy, a common renal glomerular disease. Myc1, MEK1/2 and other related kinases have been associated with one or 5 more cancers, while lamins have been implicated in muscular dystrophy and other diseases. For these reasons, siRNA directed against the genes encoding these classes of molecules would be important research and therapeutic tools.

Figure 12 illustrates four siRNAs targeting 10 different genes (Table V for 10 sequence and accession number information) that were selected according to the Formula VIII and assayed as individuals and pools in HEK293 cells. The level of siRNA induced silencing was measured using the B-DNA assay. These studies demonstrated that thirty-six out of the forty individual SMART-selected siRNA tested are functional (90%) and all 10 pools are fully functional.

15

#### **Example V. Validation of the Algorithm Using Bcl2**

Bcl-2 is a ~25kD, 205-239 amino acid, anti-apoptotic protein that contains considerable homology with other members of the BCL family including BCLX, MCL1, BAX, BAD, and BIK. The protein exists in at least two forms (Bcl2a, which 20 has a hydrophobic tail for membrane anchorage, and Bcl2b, which lacks the hydrophobic tail) and is predominantly localized to the mitochondrial membrane. While Bcl2 expression is widely distributed, particular interest has focused on the expression of this molecule in B and T cells. Bcl2 expression is down-regulated in normal germinal center B cells yet in a high percentage of follicular lymphomas, Bcl2 25 expression has been observed to be elevated. Cytological studies have identified a common translocation ((14;18)(q32;q32)) amongst a high percentage (>70%) of these lymphomas. This genetic lesion places the Bcl2 gene in juxtaposition to immunoglobulin heavy chain gene (IgH) encoding sequences and is believed to enforce inappropriate levels of gene expression, and resistance to programmed cell 30 death in the follicle center B cells. In other cases, hypomethylation of the Bcl2 promoter leads to enhanced expression and again, inhibition of apoptosis. In addition to cancer, dysregulated expression of Bcl-2 has been correlated with multiple sclerosis and various neurological diseases.

The correlation between Bcl-2 translocation and cancer makes this gene an attractive target for RNAi. Identification of siRNA directed against the bcl2 transcript (or Bcl2-IgH fusions) would further our understanding Bcl2 gene function and possibly provide a future therapeutic agent to battle diseases that result from altered 5 expression or function of this gene.

### **In Silico Identification of Functional siRNA**

To identify functional and hyperfunctional siRNA against the Bcl2 gene, the sequence for Bcl-2 was downloaded from the NCBI Unigene database and analyzed 10 using the Formula VIII algorithm. As a result of these procedures, both the sequence and SMARTscores<sup>TM</sup> of the Bcl2 siRNA were obtained and ranked according to their functionality. Subsequently, these sequences were BLAST'ed (database) to insure that the selected sequences were specific and contained minimal overlap with unrelated genes. The SMARTscores<sup>TM</sup> for the top 10 Bcl-2 siRNA are identified in **Figure 13**.

15

### **In Vivo Testing of Bcl-2 SiRNA**

Bcl-2 siRNAs having the top ten SMARTscores<sup>TM</sup> were selected and tested in a functional assay to determine silencing efficiency. To accomplish this, each of the ten duplexes were synthesized using 2'-O-ACE chemistry and transfected at 100nM 20 concentrations into cells. Twenty-four hours later assays were performed on cell extracts to assess the degree of target silencing. Controls used in these experiments included mock transfected cells, and cells that were transfected with a non-specific siRNA duplex.

25

The results of these experiments are presented below (and in **Figure 14**) and show that all ten of the selected siRNA induce 80% or better silencing of the Bcl2 message at 100nM concentrations. These data verify that the algorithm successfully identified functional Bcl2 siRNA and provide a set of functional agents that can be used in experimental and therapeutic environments.

30

|   |          |                      |                 |
|---|----------|----------------------|-----------------|
| 5 | siRNA 1  | GGGAGAUAGUGAUGAAGUA  | SEQ. ID NO. 301 |
|   | siRNA 2  | GAAGUACAUCUCAUUUAAG  | SEQ. ID NO. 302 |
|   | siRNA 3  | GUACGACAACCGGGAGAUA  | SEQ. ID NO. 303 |
|   | siRNA 4  | AGAUAGUGAUGAAGUACAU  | SEQ. ID NO. 304 |
|   | siRNA 5  | UGAAGACUCUGCUCAGUUU  | SEQ. ID NO. 305 |
|   | siRNA 6  | GCAUGCGGCCUCUGUUUGA  | SEQ. ID NO. 306 |
|   | siRNA 7  | UGCAGGCCUCUGUUUGAUUU | SEQ. ID NO. 307 |
|   | siRNA 8  | GAGAUAGUGAUGAAGUAC   | SEQ. ID NO. 308 |
|   | siRNA 9  | GGAGAUAGUGAUGAAGUAC  | SEQ. ID NO. 309 |
|   | siRNA 10 | GAAGACUCUGCUCAGUUUG  | SEQ. ID NO. 310 |

10 Bcl2 siRNA: Sense Strand, 5'→3'

#### Example VI. Sequences Selected by the Algorithm

Sequences of the siRNAs selected using Formulas (Algorithms) VIII and IX  
15 with their corresponding ranking, which have been evaluated for the silencing activity  
*in vivo* in the present study (Formula VIII and IX, respectively).

TABLE V

| Gene Name | Accession Number | SEQ. ID NO.      | FTllSeqTence        | Formula VIII | Formula IX |
|-----------|------------------|------------------|---------------------|--------------|------------|
| CLTC      | NM_004859        | SEQ. ID NO. 0301 | GAAAGAATCTGTAGAGAAA | 76           | 94.2       |
| CLTC      | NM_004859        | SEQ. ID NO. 0302 | GCAATGAGCTGTTGAAGA  | 65           | 39.9       |
| CLTC      | NM_004859        | SEQ. ID NO. 0303 | TGACAAAGGTGGATAAATT | 57           | 38.2       |
| CLTC      | NM_004859        | SEQ. ID NO. 0304 | GGAAATGGATCTCTTGAA  | 54           | 49.4       |
| CLTA      | NM_001833        | SEQ. ID NO. 0305 | GGAAAGTAATGGTCCAACA | 22           | 55.5       |
| CLTA      | NM_001833        | SEQ. ID NO. 0306 | AGACAGTTATGCAGCTATT | 4            | 22.9       |
| CLTA      | NM_001833        | SEQ. ID NO. 0307 | CCAATTCTCGGAAGCAAGA | 1            | 17         |
| CLTA      | NM_001833        | SEQ. ID NO. 0308 | GAAAGTAATGGTCCAACAG | -1           | -13        |
| CLTB      | NM_001834        | SEQ. ID NO. 0309 | GCGCCAGAGTGAACAAGTA | 17           | 57.5       |
| CLTB      | NM_001834        | SEQ. ID NO. 0310 | GAAGGTGGCCCAGCTATGT | 15           | -8.6       |
| CLTB      | NM_001834        | SEQ. ID NO. 0311 | GGAACCAGCGCCAGAGTGA | 13           | 40.5       |
| CLTB      | NM_001834        | SEQ. ID NO. 0312 | GAGCGAGATTGCAGGCATA | 20           | 61.7       |
| CALM      | U45976           | SEQ. ID NO. 0313 | GTTAGTATCTGATGACTTG | 36           | -34.6      |
| CALM      | U45976           | SEQ. ID NO. 0314 | GAAATGGAACCACTAAGAA | 33           | 46.1       |
| CALM      | U45976           | SEQ. ID NO. 0315 | GGAAATGGAACCACTAAGA | 30           | 61.2       |
| CALM      | U45976           | SEQ. ID NO. 0316 | CAACTACACTTCCAATGC  | 28           | 6.8        |
| EPS15     | NM_001981        | SEQ. ID NO. 0317 | CCACCAAGATTGATGATA  | 48           | 25.2       |

|        |           |                  |                      |     |       |
|--------|-----------|------------------|----------------------|-----|-------|
| EPS15  | NM_001981 | SEQ. ID NO. 0318 | GATCGGAACTCCAACAAGA  | 43  | 49.3  |
| EPS15  | NM_001981 | SEQ. ID NO. 0319 | AAACGGAGCTACAGATTAT  | 39  | 11.5  |
| EPS15  | NM_001981 | SEQ. ID NO. 0320 | CCACACAGCATTCTGTAA   | 33  | -23.6 |
| EPS15R | NM_021235 | SEQ. ID NO. 0321 | GAAGTTACCTTGAGCAATC  | 48  | 33    |
| EPS15R | NM_021235 | SEQ. ID NO. 0322 | GGACTTGGCCGATCCAGAA  | 27  | 33    |
| EPS15R | NM_021235 | SEQ. ID NO. 0323 | GCACTTGGATCGAGATGAG  | 20  | 1.3   |
| EPS15R | NM_021235 | SEQ. ID NO. 0324 | CAAAGACCAATTGCGTTA   | 17  | 27.7  |
| DNM2   | NM_004945 | SEQ. ID NO. 0325 | CCGAATCAATCGCATCTTC  | 6   | -29.6 |
| DNM2   | NM_004945 | SEQ. ID NO. 0326 | GACATGATCCTGCAGTTCA  | 5   | -14   |
| DNM2   | NM_004945 | SEQ. ID NO. 0327 | GAGCGAATCGTCACCACTT  | 5   | 24    |
| DNM2   | NM_004945 | SEQ. ID NO. 0328 | CCTCCGAGCTGGCGTCTAC  | -4  | -63.6 |
| ARF6   | AF93885   | SEQ. ID NO. 0329 | TCACATGGTTAACCTCTAA  | 27  | -21.1 |
| ARF6   | AF93885   | SEQ. ID NO. 0330 | GATGAGGGACGCCATAATC  | 7   | -38.4 |
| ARF6   | AF93885   | SEQ. ID NO. 0331 | CCTCTAACTACAAATCTTA  | 4   | 16.9  |
| ARF6   | AF93885   | SEQ. ID NO. 0332 | GGAAGGTGCTATCCAAAT   | 4   | 11.5  |
| RAB5A  | BC001267  | SEQ. ID NO. 0333 | GCAAGCAAGTCCTAACATT  | 40  | 25.1  |
| RAB5A  | BC001267  | SEQ. ID NO. 0334 | GGAAGAGGGAGTAGACCTTA | 17  | 50.1  |
| RAB5A  | BC001267  | SEQ. ID NO. 0335 | AGGAATCAGTGTGTAGTA   | 16  | 11.5  |
| RAB5A  | BC001267  | SEQ. ID NO. 0336 | GAAGAGGGAGTAGACCTTAC | 12  | 7     |
| RAB5B  | NM_002868 | SEQ. ID NO. 0337 | GAAAGTCAAGCCTGGTATT  | 14  | 18.1  |
| RAB5B  | NM_002868 | SEQ. ID NO. 0338 | AAAGTCAAGCCTGGTATTA  | 6   | -17.8 |
| RAB5B  | NM_002868 | SEQ. ID NO. 0339 | GCTATGAACGTGAATGATC  | 3   | -21.1 |
| RAB5B  | NM_002868 | SEQ. ID NO. 0340 | CAAGCCTGGTATTACGTTT  | -7  | -37.5 |
| RAB5C  | AF141304  | SEQ. ID NO. 0341 | GGAACAAGATCTGTCAATT  | 38  | 51.9  |
| RAB5C  | AF141304  | SEQ. ID NO. 0342 | GCAATGAACGTGAACGAAA  | 29  | 43.7  |
| RAB5C  | AF141304  | SEQ. ID NO. 0343 | CAATGAACGTGAACGAAAT  | 18  | 43.3  |
| RAB5C  | AF141304  | SEQ. ID NO. 0344 | GGACAGGAGCGGTATCACA  | 6   | 18.2  |
| EEA1   | XM_018197 | SEQ. ID NO. 0345 | AGACAGAGCTTGAGAATAA  | 67  | 64.1  |
| EEA1   | XM_018197 | SEQ. ID NO. 0346 | GAGAAGATCTTATGCAAA   | 60  | 48.7  |
| EEA1   | XM_018197 | SEQ. ID NO. 0347 | GAAGAGAAATCAGCAGATA  | 58  | 45.7  |
| EEA1   | XM_018197 | SEQ. ID NO. 0348 | GCAAGTAACTCAACTAACAA | 56  | 72.3  |
| AP2B1  | NM_001282 | SEQ. ID NO. 0349 | GAGCTAATCTGCCACATTG  | 49  | -12.4 |
| AP2B1  | NM_001282 | SEQ. ID NO. 0350 | GCAGATGAGTTACTAGAAA  | 44  | 48.9  |
| AP2B1  | NM_001282 | SEQ. ID NO. 0351 | CAACTTAATTGTCCAGAAA  | 41  | 28.2  |
| AP2B1  | NM_001282 | SEQ. ID NO. 0352 | CAACACAGGATTCTGATAAA | 33  | -5.8  |
| PLK    | NM_005030 | SEQ. ID NO. 0353 | AGATTGTGCCTAAGTCTCT  | -35 | -3.4  |
| PLK    | NM_005030 | SEQ. ID NO. 0354 | ATGAAGATCTGGAGGTGAA  | 0   | -4.3  |
| PLK    | NM_005030 | SEQ. ID NO. 0355 | TTTGAGACTTCTGCCTAA   | -5  | -27.7 |
| PLK    | NM_005030 | SEQ. ID NO. 0356 | AGATCACCCCTCCTTAAATA | 15  | 72.3  |
| GAPDH  | NM_002046 | SEQ. ID NO. 0357 | CAACGGATTGGTCGTATT   | 27  | -2.8  |

|               |           |                  |                       |     |       |
|---------------|-----------|------------------|-----------------------|-----|-------|
| GAPDH         | NM_002046 | SEQ. ID NO. 0358 | GAAATCCCATCACCATCTT   | 24  | 3.9   |
| GAPDH         | NM_002046 | SEQ. ID NO. 0359 | GACCTCAACTACATGGTTT   | 22  | -22.9 |
| GAPDH         | NM_002046 | SEQ. ID NO. 0360 | TGGTTTACATGTTCCAATA   | 9   | 9.8   |
| c-Myc         |           | SEQ. ID NO. 0361 | GAAGAAATCGATGTTGTT    | 31  | -11.7 |
| c-Myc         |           | SEQ. ID NO. 0362 | ACACAAACTTGAACAGCTA   | 22  | 51.3  |
| c-Myc         |           | SEQ. ID NO. 0363 | GGAAGAAATCGATGTTGTT   | 18  | 26    |
| c-Myc         |           | SEQ. ID NO. 0364 | GAAACGACGAGAACAGTTG   | 18  | -8.9  |
| MAP2K1        | NM_002755 | SEQ. ID NO. 0365 | GCACATGGATGGAGGTTCT   | 26  | 16    |
| MAP2K1        | NM_002755 | SEQ. ID NO. 0366 | GCAGAGAGAGCAGATTGA    | 16  | 0.4   |
| MAP2K1        | NM_002755 | SEQ. ID NO. 0367 | GAGGTTCTCTGGATCAAGT   | 14  | 15.5  |
| MAP2K1        | NM_002755 | SEQ. ID NO. 0368 | GAGCAGATTGAAGCAACT    | 14  | 18.5  |
| MAP2K2        | NM_030662 | SEQ. ID NO. 0369 | CAAAGACGATGACTTCGAA   | 37  | 26.4  |
| MAP2K2        | NM_030662 | SEQ. ID NO. 0370 | GATCAGCATTGCATGGAA    | 24  | -0.7  |
| MAP2K2        | NM_030662 | SEQ. ID NO. 0371 | TCCAGGAGTTGTCAATAA    | 17  | -4.5  |
| MAP2K2        | NM_030662 | SEQ. ID NO. 0372 | GGAAGCTGATCCACCTTGA   | 16  | 59.2  |
| KNSL1(EG5)    | NM_004523 | SEQ. ID NO. 0373 | GCAGAAATCTAAGGATATA   | 53  | 35.8  |
| KNSL1(EG5)    | NM_004523 | SEQ. ID NO. 0374 | CAACAAGGATGAAGTCTAT   | 50  | 18.3  |
| KNSL1(EG5)    | NM_004523 | SEQ. ID NO. 0375 | CAGCAGAAATCTAAGGATA   | 41  | 32.7  |
| KNSL1(EG5)    | NM_004523 | SEQ. ID NO. 0376 | CTAGATGGCTTCTCAGTA    | 39  | 3.9   |
| CyclophilinA_ | NM_021130 | SEQ. ID NO. 0377 | AGACAAGGTCCCCAAAGACA  | -16 | 58.1  |
| CyclophilinA_ | NM_021130 | SEQ. ID NO. 0378 | GGAATGGCAAGACCAAGCAA  | -6  | 36    |
| CyclophilinA_ | NM_021130 | SEQ. ID NO. 0379 | AGAATTATTCCAGGGTTA    | -3  | 16.1  |
| CyclophilinA_ | NM_021130 | SEQ. ID NO. 0380 | GCAGACAAAGGTCCCCAAAGA | 8   | 8.9   |
| LAMIN A/C     | NM_170707 | SEQ. ID NO. 0381 | AGAAGCAGCTTCAGGATGA   | 31  | 38.8  |
| LAMIN A/C     | NM_170707 | SEQ. ID NO. 0382 | GAGCTTGACTTCCAGAAGA   | 33  | 22.4  |
| LAMIN A/C     | NM_170707 | SEQ. ID NO. 0383 | CCACCGAAGTTCACCCCTAA  | 21  | 27.5  |
| LAMIN A/C     | NM_170707 | SEQ. ID NO. 0384 | GAGAAGAGCTCCTCCATCA   | 55  | 30.1  |
| CyclophilinB  | M60857    | SEQ. ID NO. 0385 | GAAAGAGCATCTACGGTGA   | 41  | 83.9  |
| CyclophilinB  | M60857    | SEQ. ID NO. 0386 | GAAAGGATTGGCTACAAA    | 53  | 59.1  |
| CyclophilinB  | M60857    | SEQ. ID NO. 0387 | ACAGCAAATTCCATCGTGT   | -20 | 28.8  |
| CyclophilinB  | M60857    | SEQ. ID NO. 0388 | GGAAAGACTGTTCCAAAAAA  | 2   | 27    |
| DBI1          | NM_020548 | SEQ. ID NO. 0389 | CAACACGCCATCCTCTA     | 27  | -7.6  |
| DBI2          | NM_020548 | SEQ. ID NO. 0390 | CATGAAAGCTTACATCAAC   | 25  | -30.8 |
| DBI3          | NM_020548 | SEQ. ID NO. 0391 | AAGATGCCATGAAAGCTTA   | 17  | 22    |
| DBI4          | NM_020548 | SEQ. ID NO. 0392 | GCACATAACGCCCTGAGTCT  | 15  | 3.9   |
| rLUC1         |           | SEQ. ID NO. 0393 | GATCAAATCTGAAGAAGGA   | 57  | 49.2  |
| rLUC2         |           | SEQ. ID NO. 0394 | GCCAAGAAGTTCTAATA     | 50  | 13.7  |
| rLUC3         |           | SEQ. ID NO. 0395 | CAGCATATCTGAACCATT    | 41  | -2.2  |
| rLUC4         |           | SEQ. ID NO. 0396 | GAACAAAGGAAACGGATGA   | 39  | 29.2  |
| SeAP1         | NM_031313 | SEQ. ID NO. 0397 | CGGAAACGGTCCAGGCTAT   | 6   | 26.9  |

|          |             |                  |                     |     |       |
|----------|-------------|------------------|---------------------|-----|-------|
| SeAP2    | NM_031313   | SEQ. ID NO. 0398 | GCTTCGAGCAGACATGATA | 4   | -11.2 |
| SeAP3    | NM_031313   | SEQ. ID NO. 0399 | CCTACACGGTCCTCCTATA | 4   | 4.9   |
| SeAP4    | NM_031313   | SEQ. ID NO. 0400 | GCCAAGAACCTCATCATCT | 1   | -9.9  |
| fLUC1    |             | SEQ. ID NO. 0401 | GATATGGCTGAATACAAA  | 54  | 40.4  |
| fLUC2    |             | SEQ. ID NO. 0402 | GCACTCTGATTGACAAATA | 47  | 54.7  |
| fLUC3    |             | SEQ. ID NO. 0403 | TGAAGTCTCTGATTAAGTA | 46  | 34.5  |
| fLUC4    |             | SEQ. ID NO. 0404 | TCAGAGAGATCCTCATAAA | 40  | 11.4  |
| mCyclo_1 | NM_008907   | SEQ. ID NO. 0405 | GCAAGAAGATCACCATTTC | 52  | 46.4  |
| mCyclo_2 | NM_008907   | SEQ. ID NO. 0406 | GAGAGAAATTGAGGATGA  | 36  | 70.7  |
| mCyclo_3 | NM_008907   | SEQ. ID NO. 0407 | GAAAGGATTTGGCTATAAG | 35  | -1.5  |
| mCyclo_4 | NM_008907   | SEQ. ID NO. 0408 | GAAAGAAGGCATGAACATT | 27  | 10.3  |
| BCL2_1   | NM_000633   | SEQ. ID NO. 0409 | GGGAGATAGTGATGAAGTA | 21  | 72    |
| BCL2_2   | NM_000633   | SEQ. ID NO. 0410 | GAAGTACATCCATTATAAG | 1   | 3.3   |
| BCL2_3   | NM_000633   | SEQ. ID NO. 0411 | GTACGACAACCGGGAGATA | 1   | 35.9  |
| BCL2_4   | NM_000633   | SEQ. ID NO. 0412 | AGATAGTGATGAAGTACAT | -12 | 22.1  |
| BCL2_5   | NM_000633   | SEQ. ID NO. 0413 | TGAAGACTCTGCTCAGTT  | 36  | 19.1  |
| BCL2_6   | NM_000633   | SEQ. ID NO. 0414 | GCATGCGGCCTCTGTTGA  | 5   | -9.7  |
| QB1      | NM_003365.1 | SEQ. ID NO. 0415 | GCACACAGCUUACUACAU  | 52  | -4.8  |
| QB2      | NM_003365.1 | SEQ. ID NO. 0416 | GAAAUGCCCUGGUACUCA  | 49  | 22.1  |
| QB3      | NM_003365.1 | SEQ. ID NO. 0417 | GAAGGAACGUGAUGUGAU  | 34  | 22.9  |
| QB4      | NM_003365.1 | SEQ. ID NO. 0418 | GCACUACUCCUGUGUGUGA | 28  | 20.4  |
| ATE1-1   | NM_007041   | SEQ. ID NO. 0419 | GAACCCAGCUGGAGAACUU | 45  | 15.5  |
| ATE1-2   | NM_007041   | SEQ. ID NO. 0420 | GAUUAUACAGUGUGAUUUA | 40  | 12.2  |
| ATE1-3   | NM_007041   | SEQ. ID NO. 0421 | GUACUACGAUCCUGAUUAU | 37  | 32.9  |
| ATE1-4   | NM_007041   | SEQ. ID NO. 0422 | GUGCCGACCUUACAAUUU  | 35  | 18.2  |
| EGFR-1   | NM_005228   | SEQ. ID NO. 0423 | GAAGGAAACTGAATTCAA  | 68  | 79.4  |
| EGFR-1   | NM_005228   | SEQ. ID NO. 0424 | GGAAATATGTACTACGAA  | 49  | 49.5  |
| EGFR-1   | NM_005228   | SEQ. ID NO. 0425 | CCACAAAGCAGTGAATT   | 41  | 7.6   |
| EGFR-1   | NM_005228   | SEQ. ID NO. 0426 | GTAACAAAGCTCACGCAGT | 40  | 25.9  |

### Example VII. Genome-Wide Application of the Algorithm

The examples described above demonstrate that the algorithm(s) can successfully identify functional siRNA and that these duplexes can be used to induce the desirable phenotype of transcriptional knockdown or knockout. Each gene or family of genes in each organism plays an important role in maintaining physiological homeostasis and the algorithm can be used to develop functional, highly functional, or hyperfunctional siRNA to each gene. To accomplish this for the human genome, the entire online ncbi refseq database was accessed through Entrez (efetch). The database

was processed through Formula VIII. For each gene the top 80–100 scores for siRNAs were obtained and BLAST'ed to insure that the selected sequences are specific in targeting the gene of choice. These sequences are provided on the enclosed CDs in electronic form. Accordingly, Applicants hereby incorporate by 5 reference the material submitted herewith, in duplicate on the compact disks labeled COPY 1 – TABLES PART, DISK 1/1, TABLES 12 –15, Filed with RO/US under PCT AI sec. 801(a), Operating System: MS-Windows, COPY 2 – TABLES PART, DISK 1/1, TABLES 12 –15, Filed with RO/US under PCT AI sec. 801(a), Operating System: MS-Windows, COPY 3 – TABLES PART, DISK 1/1, TABLES 12 –15, 10 Filed with RO/US under PCT AI sec. 801(a), Operating System: MS-Windows,; which copies are identical, in files entitled Table\_12.txt, date of creation June 26, 2003, with a size of 31,045 kb; Table\_13.txt; date of creation November 13, 2003, with a size of 78,451 kb; Table\_14.txt, date of creation November 13, 2003, with a size of 454 kb; and Table\_15.txt date of creation November 13, 2003, with a size of 15 1,690 kb.

With respect to the disks, there are four tables on each disk copy in text format: Tables XII –XV. Table XII, which is located in a file entitled Table\_12.txt, provides a list of the 80-100 sequences for each target, identified by Formula VIII as 20 having the highest relative SMARTscores<sup>TM</sup> for the target analyzed. Table XIII, which is located in a file entitled Table\_13.txt, provides the SMARTscores<sup>TM</sup>, and for each gene, a pool pick of up to four sequences is denoted. (The denotation of “1” in Table XIII means that it is a pool pick.) These pool pick sequences represent the most functional siRNAs for the corresponding target. Any 1, 2, 3, or 4 of the pool pick 25 sequences could be used for gene silencing. Further, sequences that are not denoted as pool pick sequences, but that are included on the compact disks may also be used for gene silencing either alone or in combination with other sequences. However, their individual relative functionality would be less than that of a pool pick sequence. Table XIV, which is located in a file entitled Table\_14.txt, provides an identification 30 of genes by accession number, and Table XV, which is located in a file entitled Table\_15.txt, provides a short name for the genes identified on the disk. The information contained on the disks is part of this patent application and are incorporated into the specification by reference. One may use these tables in order to identify functional siRNAs for the gene provided therein, by simply looking for the

gene of interest and an siRNA that is listed as functional. Preferably, one would select one or more of the siRNA that most optimized for the target of interest and is denoted as a pool pick.

5

**Table XII: siRNA selected by Formula VIII**

See data submitted herewith on a CD-ROM in accordance with PCT Administrative Instructions Section 801(a)

10

**Table XIII: SMARTscores™**

See data submitted herewith on a CD-ROM in accordance with PCT Administrative Instructions Section 801(a)

15

**Table XIV: Identification of Targets**

See data submitted herewith on a CD-ROM in accordance with PCT

Administrative Instructions Section 801(a)

20

**Table XV: Description of Targets**

See data submitted herewith on a CD-ROM in accordance with PCT Administrative Instructions Section 801(a)

25

Many of the genes to which the described siRNA are directed play critical roles in disease etiology. For this reason, the siRNA listed in the accompanying compact disk may potentially act as therapeutic agents. A number of prophetic examples follow and should be understood in view of the siRNA that are identified on the accompanying CD. To isolate these siRNA, the appropriate message sequence for each gene is analyzed using one of the before mentioned formulas (preferably formula VIII) to identify potential siRNA targets. Subsequently these targets are BLAST'ed to eliminate homology with potentially off-targets.

30

The list of potential disease targets is extensive. For instance, over-expression of Bcl10 has been implicated in the development of MALT lymphoma (mucosa associated lymphoid tissue lymphoma) and thus, functional, highly functional, or hyperfunctional siRNA directed against that gene (e.g. SEQ. ID NO. 0427: GGAAACCUCAUUGCUCUAA; SEQ. ID NO. 0428:

GAAAGAACCUUGCCGAUCA; SEQ. ID NO. 0429;  
GGAAAUAACAUCAUCAGAGCUUA, or SEQ. ID NO. 0430:  
GAAAGUAUGUGUCUUAAGU) may contribute to treatment of this disorder.

5 In another example, studies have shown that molecules that inhibit glutamine:fructose-6-phosphate aminotransferase (GFA) may act to limit the symptoms suffered by Type II diabetics. Thus, functional, highly functional, or hyperfunctional siRNA directed against GFA (also known as GFPT1: siRNA = SEQ. ID NO. 0433 UGAAACGGCUGCCUGAUUU; SEQ. ID NO. 0434

10 GAAGUUACCUCUUACAUUU; SEQ. ID NO. 0435  
GUACGAAACUGUAUGAUUA; SEQ. ID NO. 0436  
GGACGAGGCUAUCAUUAUG) may contribute to treatment of this disorder.

In another example, the von Hippel-Lindau (VHL) tumor suppressor has been observed to be inactivated at a high frequency in sporadic clear cell renal cell carcinoma (RCC) and RCCs associated with VHL disease. The VHL tumor suppressor targets hypoxia-inducible factor-1 alpha (HIF-1 alpha), a transcription factor that can induce vascular endothelial growth factor (VEGF) expression, for ubiquitination and degradation. Inactivation of VHL can lead to increased levels of HIF-1 alpha, and subsequent VEGF over expression. Such over expression of VEGF has been used to explain the increased (and possibly necessary) vascularity observed in RCC. Thus, functional, highly functional, or hyperfunctional siRNAs directed against either HIF-1 alpha (SEQ. ID NO. 0437 GAAGGAACCUGAUGCUUUA; SEQ. ID NO. 0438 GCAUUAUCUAGAAGGUAU; SEQ. ID NO. 0439

25 GAACAAAUAACAUGGGAUUA; SEQ. ID NO. 0440  
GGACACAGAUUUAGACUUG) or VEGF (SEQ. ID NO. 0441  
GAACGUACUUGCAGAUGUG; SEQ. ID NO. 0442  
GAGAAAGCAUUUGUUUGUA; SEQ. ID NO. 0443  
GGAGAAAGCAUUUGUUUGU; SEQ. ID NO. 0444

30 CGAGGCAGCUUGAGUUAAA) may be useful in the treatment of renal cell carcinoma.

In another example, gene expression of platelet derived growth factor A and B (PDGF-A and PDGF-B) has been observed to be increased 22- and 6-fold,

respectively, in renal tissues taken from patients with diabetic nephropathy as compared with controls. These findings suggest that over expression of PDGF A and B may play a role in the development of the progressive fibrosis that characterizes human diabetic kidney disease. Thus, functional, highly functional, or hyperfunctional

5 siRNAs directed against either PDGF A

(SEQ. ID NO. 0445: GGUAAGAUUAUUGUGCUUUA;  
SEQ. ID NO. 0446: CCGCAAAUAUGCAGAAUUA;  
SEQ. ID NO. 0447: GGAUGUACAUUUGGCGUGUUA;  
SEQ. ID NO. 0448: GGUGAAGUUUUGUAUGUUUA) or

10

PDGF B

(SEQ. ID NO. 0449: CCGAGGAGCUUUAUGAGAU;  
SEQ. ID NO. 0450: GCUCCGCGCUUUCGAUUU;  
SEQ. ID NO. 0451 GAGCAGGAAUUGGUGAGAUG;  
15 SEQ. ID NO. 0452: GAACUUGGGAUAAAGAGUGU;  
SEQ. ID NO. 0453 CCGAGGAGCUUUAUGAGAU;  
SEQ. ID NO. 0454 UUUAUGAGAUGCUGAGUGA) may be useful in the treatment of this form of kidney disorder.

20

In another example, a strong correlation exists between the over-expression of glucose transporters (*e.g.* GLUT12) and cancer cells. It is predicted that cells undergoing uncontrolled cell growth up-regulate GLUT molecules so that they can cope with the heightened energy needs associated with increased rates of proliferation and metastasis. Thus, siRNA-based therapies that target the molecules such as

25 GLUT1 (also known as SLC2A1: siRNA=

SEQ. ID NO.: 0455 GCAAUGAUGUCCAGAAGAA;  
SEQ. ID NO.: 0456 GAAGAAUUAUUCAGGACUUA;  
SEQ. ID NO.: 0457 GAAGAGAGUCGGCAGAUGA;  
SEQ. ID NO.: 0458 CCAAGAGUGUGCUAAAGAA)

30

GLUT12 (also known as SLCA12: siRNA =

SEQ. ID NO. 0459: GAGACACUCUGAAAUGAUA;  
SEQ. ID NO. 0460: GAAAUGAUGUGGAUAAGAG;  
SEQ. ID NO. 0461: GAUCAAAUCCUCCCCUGAAA;

SEQ. ID NO. 0462: UGAAUGAGCUGAUGAUUGU) and other related transporters, may be of value in treating a multitude of malignancies.

5 The siRNA sequences listed above are presented in a 5' → 3' sense strand direction. In addition, siRNA directed against the targets listed above as well as those directed against other targets and listed in the accompanying compact disk may be useful as therapeutic agents.

#### **Example VIII. Evidence for the Benefits of Pooling**

10 Evidence for the benefits of pooling have been demonstrated using the reporter gene, luciferase. Ninety siRNA duplexes were synthesized using Dharmacon proprietary ACE® chemistry against one of the standard reporter genes: firefly luciferase. The duplexes were designed to start two base pairs apart and to cover approximately 180 base pairs of the luciferase gene (see sequences in **Table III**).  
15 Subsequently, the siRNA duplexes were co-transfected with a luciferase expression reporter plasmid into HEK293 cells using standard transfection protocols and luciferase activity was assayed at 24 and 48 hours.

20 Transfection of individual siRNAs showed standard distribution of inhibitory effect. Some duplexes were active, while others were not. **Figure 15** represents a typical screen of ninety siRNA duplexes (SEQ. ID NO. 0032- 0120) positioned two base pairs apart. As the figure suggests, the functionality of the siRNA duplex is determined more by a particular sequence of the oligonucleotide than by the relative oligonucleotide position within a gene or excessively sensitive part of the mRNA,  
25 which is important for traditional anti-sense technology.

When two continuous oligonucleotides were pooled together, a significant increase in gene silencing activity was observed. (See **Figure 16**) A gradual increase in efficacy and the frequency of pools functionality was observed when the number of 30 siRNAs increased to 3 and 4. (**Figures 16, 17**). Further, the relative positioning of the oligonucleotides within a pool did not determine whether a particular pool was functional (see **Figure 18**, in which 100% of pools of oligonucleotides distanced by 2, 10 and 20 base pairs were functional).

However, relative positioning may nonetheless have an impact. An increased functionality may exist when the siRNA are positioned continuously head to toe (5' end of one directly adjacent to the 3' end of the others).

5        Additionally, siRNA pools that were tested performed at least as well as the best oligonucleotide in the pool, under the experimental conditions whose results are depicted in **Figure 19**. Moreover, when previously identified non-functional and marginally (semi) functional siRNA duplexes were pooled together in groups of five at a time, a significant functional cooperative action was observed. (See **Figure 20**)

10      In fact, pools of semi-active oligonucleotides were 5 to 25 times more functional than the most potent oligonucleotide in the pool. Therefore, pooling several siRNA duplexes together does not interfere with the functionality of the most potent siRNAs within a pool, and pooling provides an unexpected significant increase in overall functionality

15

#### **Example IX. Pooling Across Species**

Experiments were performed on the following genes:  $\beta$ -galactosidase, Renilla luciferase, and Secreted alkaline phosphatase, which demonstrates the benefits of pooling. (see **Figure 21**) Approximately 50% of individual siRNAs designed to silence the above-specified genes were functional, while 100% of the pools that contain the same siRNA duplexes were functional.

#### **Example X. Highly Functional siRNA**

25      Pools of five siRNAs in which each two siRNAs overlap to 10-90% resulted in 98% functional entities (>80% silencing). Pools of siRNAs distributed throughout the mRNA that were evenly spaced, covering an approximate 20 – 2000 base pair range, were also functional. When the pools of siRNA were positioned continuously head to tail relative to mRNA sequences and mimicked the natural products of Dicer cleaved long double stranded RNA, 98% of the pools evidenced highly functional activity (>95% silencing).

#### **Example XI. Human cyclophyline**

**Table III** above lists the siRNA sequences for the human cyclophyline protein. A particularly functional siRNA may be selected by applying these sequences to any of Formula I to VII above.

5        Alternatively, one could pool 2, 3, 4, 5 or more of these sequences to create a kit for silencing a gene. Preferably, within the kit there would be at least one sequence that has a relatively high predicted functionality when any of Formulas I - VII is applied.

10      **Example XII. Sample Pools of siRNAs and Their Application to Human Disease**

The genetic basis behind human disease is well documented and siRNA may be used as both research or diagnostic tools and therapeutic agents; either individually or in pools. Genes involved in signal transduction, the immune response, apoptosis, DNA repair, cell cycle control, and a variety of other physiological functions have 15 clinical relevance and therapeutic agents that can modulate expression of these genes may alleviate some or all of the associated symptoms. In some instances, these genes can be described as a member of a family or class of genes and siRNA (randomly, conventionally, or rationally designed) can be directed against one or multiple members of the family to induce a desired result.

20        To identify rationally designed siRNA to each gene, the sequence was analyzed using Formula VIII to identify a SMARTpool containing the functional sequences. To confirm the activity of these sequences, the siRNA are introduced into a cell type of choice (e.g. HeLa cells, HEK293 cells) and the levels of the appropriate 25 message are analyzed using one of several art proven techniques. SiRNA having heightened levels of potency can be identified by testing each of the before mentioned duplexes at increasingly limiting concentrations. Similarly, siRNA having increased levels of longevity can be identified by introducing each duplex into cells and testing functionality at 24, 48, 72, 96, 120, 144, 168, and 192 hours after transfection. Agents 30 that induce >95% silencing at sub-nanomolar concentrations and/or induce functional levels of silencing for >96 hours are considered hyperfunctional.

The following are non-limiting examples of families of proteins to which siRNA described in this document are targeted against:

### Transporters, Pumps, and Channels

5 Transporters, pumps, and channels represent one class of genes that are attractive targets for siRNAs. One major class of transporter molecules are the ATP-binding cassette (ABC) transporters. To date, nearly 50 human ABC-transporter genes have been characterized and have been shown to be involved in a variety of physiological functions including transport of bile salts, nucleosides, chloride ions, 10 cholesterol, toxins, and more. Predominant among this group are MDR1 (which encodes the P-glycoprotein, NP\_000918), the MDR-related proteins (MRP1-7), and the breast cancer resistance protein (BCRP). In general, these transporters share a common structure, with each protein containing a pair of ATP-binding domains (also known as nucleotide binding folds, NBF) and two sets of transmembrane (TM) 15 domains, each of which typically contains six membrane-spanning  $\alpha$ -helices. The genes encoding this class of transporter are organized as either full transporters (*i.e.* containing two TM and two NBF domains) or as half transporters that assemble as either homodimers or heterodimers to create functional transporters. As a whole, members of the family are widely dispersed throughout the genome and show a high 20 degree of amino acid sequence identity among eukaryotes.

ABC-transporters have been implicated in several human diseases. For instance, molecular efflux pumps of this type play a major role in the development of drug resistance exhibited by a variety of cancers and pathogenic microorganisms. In 25 the case of human cancers, increased expression of the MDR1 gene and related pumps have been observed to generate drug resistance to a broad collection of commonly used chemotherapeutics including doxorubicin, daunorubicin, vinblastine, vincristine, colchicines. In addition to the contribution these transporters make to the development of multi-drug resistance, there are currently 13 human genetic diseases 30 associated with defects in 14 different transporters. The most common of these conditions include cystic fibrosis, Stargardt disease, age-related macular degeneration, adrenoleukodystrophy, Tangier disease, Dubin-Johnson syndrome and progressive familial intrahepatic cholestasis. For this reason, siRNAs directed against members of this, and related, families are potentially valuable research and therapeutic tools.

With respect to channels, analysis of *Drosophila* mutants has enabled the initial molecular isolation and characterization of several distinct channels including (but not limited to) potassium (K<sup>+</sup>) channels. This list includes shaker (Sh), which encodes a voltage activated K<sup>+</sup> channel, slowpoke (Slo), a Ca<sup>2+</sup> activated K<sup>+</sup> channel, 5 and ether-a-go-go (Eag). The Eag family is further divided into three subfamilies: Eag, Elk (eag-like K channels), and Erg (Eag related genes).

The Erg subfamily contains three separate family members (Erg1-3) that are 10 distantly related to the sh family of voltage activated K<sup>+</sup> channels. Like sh, erg polypeptides contain the classic six membrane spanning architecture of K<sup>+</sup> channels (S1-S6) but differ in that each includes a segment associated with the C-terminal cytoplasmic region that is homologous to cyclic nucleotide binding domains (cNBD). Like many isolated ion channel mutants, erg mutants are temperature-sensitive 15 paralytics, a phenotype caused by spontaneous repetitive firing (hyperactivity) in neurons and enhanced transmitter release at the neuromuscular junction.

Initial studies on the tissue distribution of all three members of the erg subfamily show two general patterns of expression. Erg1 and erg3 are broadly 20 expressed throughout the nervous system and are observed in the heart, the superior mesenteric ganglia, the celiac ganglia, the retina, and the brain. In contrast, erg2 shows a much more restricted pattern of expression and is only observed in celiac ganglia and superior mesenteric ganglia. Similarly, the kinetic properties of the three erg potassium channels are not homogeneous. Erg1 and erg2 channels are relatively slow activating delayed rectifiers whereas the erg3 current activates rapidly and then 25 exhibits a predominantly transient component that decays to a sustained plateau. The current properties of all three channels are sensitive to methanesulfonanilides, suggesting a high degree of conservation in the pore structure of all three proteins.

Recently, the erg family of K<sup>+</sup> channels has been implicated in human disease. 30 Consistent with the observation that erg1 is expressed in the heart, single strand conformation polymorphism and DNA sequence analyses have identified HERG (human erg1) mutations in six long-QT-syndrome (LQT) families, an inherited disorder that results in sudden death from a ventricular tachyarrhythmia. Thus siRNA

directed against this group of molecules (e.g. KCNH1-8) will be of extreme therapeutic value.

Another group of channels that are potential targets of siRNAs are 5 the CLCA family that mediate a  $\text{Ca}^{2+}$ -activated  $\text{Cl}^-$  conductance in a variety of tissues. To date, two bovine (bCLC1; bCLCA2 (Lu-ECAM-1)), three mouse (mCLCA1; mCLCA2; mCLCA3) and four human (hCLCA1; hCLCA2; hCLCA3; hCLCA4) CLCA family members have been isolated and patch-clamp studies with transfected human embryonic kidney (HEK-293) cells have shown that bCLCA1, 10 mCLCA1, and hCLCA1 mediate a  $\text{Ca}^{2+}$ -activated  $\text{Cl}^-$  conductance that can be inhibited by the anion channel blocker DIDS and the reducing agent dithiothreitol (DTT).

The protein size, structure, and processing seem to be similar among different 15 CLCA family members and has been studied in greatest detail for Lu-ECAM-1. The Lu-ECAM-1 open reading frame encodes a precursor glycoprotein of 130 kDa that is processed to a 90-kDa amino-terminal cleavage product and a group of 30- to 40-kDa glycoproteins that are glycosylation variants of a single polypeptide derived from its carboxy terminus. Both subunits are associated with the outer cell surface, but only 20 the 90-kDa subunit is thought to be anchored to the cell membrane via four transmembrane domains.

Although the protein processing and function appear to be conserved among 25 CLCA homologs, significant differences exist in their tissue expression patterns. For example, bovine Lu-ECAM-1 is expressed primarily in vascular endothelia, bCLCA1 is exclusively detected in the trachea, and hCLCA1 is selectively expressed in a subset of human intestinal epithelial cells. Thus the emerging picture is that of a multigene family with members that are highly tissue specific, similar to the ClC family of voltage-gated  $\text{Cl}^-$  channels. The human channel, hCLCA2, is particular 30 interesting from a medical and pharmacological standpoint. CLCA2 is expressed on the luminal surface of lung vascular endothelia and serves as an adhesion molecule for lung metastatic cancer cells, thus mediating vascular arrest and lung colonization. Expression of this molecule in normal mammary epithelium is consistently lost in human breast cancer and in nearly all tumorigenic breast cancer cell lines. Moreover,

re-expression of hCLCA2 in human breast cancer cells abrogates tumorigenicity in nude mice, implying that hCLCA2 acts as a tumour suppressor in breast cancer. For these reasons, siRNA directed against CLCA family members and related channels may prove to be valuable in research and therapeutic venues.

5

### **Transporters Involved in Synaptic Transmission.**

Synaptic transmission involves the release of a neurotransmitter into the synaptic cleft, interaction of that transmitter with a postsynaptic receptor, and subsequent removal of the transmitter from the cleft. In most synapses the signal is 10 terminated by a rapid reaccumulation of the neurotransmitter into presynaptic terminals. This process is catalyzed by specific neurotransmitter transporters that are often energized by the electrochemical gradient of sodium across the plasma membrane of the presynaptic cells.

15

Aminobutyric acid (GABA) is the major inhibitory neurotransmitter in the central nervous system. The inhibitory action of GABA, mediated through GABA<sub>A</sub>/GABA<sub>B</sub> receptors, and is regulated by GABA transporters (GATs), integral membrane proteins located perisynaptically on neurons and glia. So far four different carriers (GAT1-GAT4) have been cloned and their cellular distribution has been 20 partly worked out. Comparative sequence analysis has revealed that GABA transporters are related to several other proteins involved in neurotransmitter uptake including gamma-aminobutyric acid transporters, monoamine transporters, amino acid transporters, certain "orphan" transporters, and the recently discovered bacterial transporters. Each of these proteins has a similar 12 transmembrane helices topology 25 and relies upon the Na<sup>+</sup>/Cl<sup>-</sup> gradient for transport function. Transport rates are dependent on substrate concentrations, with half-maximal effective concentrations for transport frequently occurring in the submicromolar to low micromolar range. In addition, transporter function is bidirectional, and non-vesicular efflux of transmitter may contribute to ambient extracellular transmitter levels.

30

Recent evidence suggests that GABA transporters, and neurotransmitter transporters in general, are not passive players in regulating neuronal signaling; rather, transporter function can be altered by a variety of initiating factors and signal transduction cascades. In general, this functional regulation occurs in two ways,

either by changing the rate of transmitter flux through the transporter or by changing the number of functional transporters on the plasma membrane. A recurring theme in transporter regulation is the rapid redistribution of the transporter protein between intracellular locations and the cell surface. In general, this functional modulation 5 occurs in part through activation of second messengers such as kinases, phosphatases, arachidonic acid, and pH. However, the mechanisms underlying transporter phosphorylation and transporter redistribution have yet to be fully elucidated.

GABA transporters play a pathophysiological role in a number of human 10 diseases including temporal lobe epilepsy and are the targets of pharmacological interventions. Studies in seizure sensitive animals show some (but not all) of the GAT transporters have altered levels of expression at times prior to and post seizure, suggesting this class of transporter may affect epileptogenesis, and that alterations following seizure may be compensatory responses to modulate seizure activity. For 15 these reasons, siRNAs directed against members of this family of genes (including but not limited to SLCG6A1-12) may prove to be valuable research and therapeutic tools.

### **Organic Ion Transporters.**

The human body is continuously exposed to a great variety of xenobiotics, via 20 food, drugs, occupation, and environment. Excretory organs such as kidney, liver, and intestine defend the body against the potentially harmful effects of these compounds by transforming them into less active metabolites that are subsequently secreted from the system.

25 Carrier-mediated transport of xenobiotics and their metabolites exist for the active secretion of organic anions and cations. Both systems are characterized by a high clearance capacity and tremendous diversity of substances accepted, properties that result from the existence of multiple transporters with overlapping substrate specificities. The class of organic anion transporters plays a critical role in the 30 elimination of a large number of drugs (*e.g.*, antibiotics, chemotherapeutics, diuretics, nonsteroidal anti-inflammatory drugs, radiocontrast agents, cytostatics); drug metabolites (especially conjugation products with glutathione, glucuronide, glycine, sulfate, acetate); and toxicants and their metabolites (*e.g.*, mycotoxins, herbicides,

plasticizers, glutathione *S*-conjugates of polyhaloalkanes, polyhaloalkenes, hydroquinones, aminophenols), many of which are specifically harmful to the kidney.

Over the past couple of years the number of identified anion transporting molecules has grown tremendously. Uptake of organic anions ( $OA^-$ ) across the basolateral membrane is mediated by the classic sodium-dependent organic anion transport system, which includes  $\alpha$ -ketoglutarate ( $\alpha$ -KG $^{2-}$ )/ $OA^-$  exchange via the organic anion transporter (OAT1) and sodium-ketoglutarate cotransport via the  $Na^+$ /dicarboxylate cotransporter (SDCT2). The organic anion transporting polypeptide, Oatp1, and the kidney-specific OAT-K1 and OAT-K2 are seen as potential molecules that mediate facilitated  $OA^-$  efflux but could also be involved in reabsorption via an exchange mechanism. Lastly the PEPT1 and PEPT2 mediate luminal uptake of peptide drugs, whereas CNT1 and CNT2 are involved in reabsorption of nucleosides

The organic anion-transporting polypeptide 1 (Oatp1) is a  $Na^+$ - and ATP-independent transporter originally cloned from rat liver. The tissue distribution and transport properties of the Oatp1 gene product are complex. Oatp1 is localized to the basolateral membrane of hepatocytes, and is found on the apical membrane of S3 proximal tubules. Studies with transiently transfected cells (e.g. HeLa cells) have indicated that Oatp1 mediates transport of a variety of molecules including taurocholate, estrone-3-sulfate, aldosterone, cortisol, and others. The observed uptake of taurocholate by Oatp1 expressed in *X. laevis* oocytes is accompanied by efflux of GSH, suggesting that transport by this molecule may be glutathione dependent.

Computer modeling suggests that members of the Oatp family are highly conserved, hydrophobic, and have 12 transmembrane domains. Decreases in expression of Oatp family members have been associated with cholestatic liver diseases and human hepatoblastomas, making this family of proteins of key interest to researchers and the medical community. For these reasons, siRNAs directed against OAT family members (including but not limited to SLC21A2, 3, 6, 8, 9, 11, 12, 14, 15, and related transporters) are potentially useful as research and therapeutic tools.

### **Nucleoside transporters.**

Nucleoside transporters play key roles in physiology and pharmacology.

Uptake of exogenous nucleosides is a critical first step of nucleotide synthesis in tissues such as bone marrow and intestinal epithelium and certain parasitic organisms that lack *de novo* pathways for purine biosynthesis. Nucleoside transporters also control the extracellular concentration of adenosine in the vicinity of its cell surface receptors and regulate processes such as neurotransmission and cardiovascular activity. Adenosine itself is used clinically to treat cardiac arrhythmias, and nucleoside transport inhibitors such as dipyridamole, dilazep, and drafazine function as coronary vasodilators.

10

In mammals, plasma membrane transport of nucleosides is brought about by members of the concentrative,  $\text{Na}^+$ -dependent (CNT) and equilibrative,  $\text{Na}^+$ -independent (ENT) nucleoside transporter families. CNTs are expressed in a tissue-specific fashion; ENTs are present in most, possibly all, cell types and are responsible for the movement of hydrophilic nucleosides and nucleoside analogs down their concentration gradients. In addition, structure/function studies of ENT family members have predicted these molecules to contain eleven transmembrane helical segments with an amino terminus that is intracellular and a carboxyl terminus that is extracellular. The proteins have a large glycosylated loop between TMs 1 and 2 and a large cytoplasmic loop between TMs 6 and 7. Recent investigations have implicated the TM 3-6 region as playing a central role in solute recognition. The medical importance of the ENT family of proteins is broad. In humans adenosine exerts a range of cardioprotective effects and inhibitors of ENTs are seen as being valuable in alleviating a variety of cardio/cardiovascular ailments. In addition, responses to nucleoside analog drugs has been observed to vary considerably amongst *e.g.* cancer patients. While some forms of drug resistance have been shown to be tied to the up-regulation of ABC-transporters (*e.g.* MDR1), resistance may also be the result of reduced drug uptake (*i.e.* reduced ENT expression). Thus, a clearer understanding of ENT transporters may aid in optimizing drug treatments for patients suffering a wide range of malignancies. For these reasons, siRNAs directed against this class of molecules (including SLC28A1-3, SLC29A1-4, and related molecules) may be useful as therapeutic and research tools.

### Sulfate Transporters

All cells require inorganic sulfate for normal function. Sulfate is the fourth most abundant anion in human plasma and is the major source of sulfur in many organisms. Sulfation of extracellular matrix proteins is critical for maintaining normal cartilage metabolism and sulfate is an important constituent of myelin  
5 membranes found in the brain

Because sulfate is a hydrophilic anion that cannot passively cross the lipid bilayer of cell membranes, all cells require a mechanism for sulfate influx and efflux to ensure an optimal supply. To date, a variety of sulfate transporters have been  
10 identified in tissues from many origins. These include the renal sulfate transporters (NaSi-1 and Sat-1), the ubiquitously expressed diastrophic dysplasia sulfate transporter (DTDST), the intestinal sulfate transporter (DRA), and the erythrocyte anion exchanger (AE1). Most, if not all, of these molecules contain the classic 12 transmembrane spanning domain architecture commonly found amongst members of  
15 the anion transporter superfamily.

Recently three different sulfate transporters have been associated with specific human genetic diseases. Family members SLC26A2, SLC26A3, and SLC26A4 have been recognized as the disease genes mutated in diastrophic dysplasia, congenital  
20 chloride diarrhea (CLD), and Pendred syndrome (PDS), respectively. DTDST is a particularly complex disorder. The gene encoding this molecule maps to chromosome 5q, and encodes two distinct transcripts due to alternative exon usage. In contrast to other sulfate transporters (e.g. Sat-1) anion movement by the DTDST protein is markedly inhibited by either extracellular chloride or bicarbonate. Impaired function  
25 of the DTDST gene product leads to undersulfation of proteoglycans and a complex family of recessively inherited osteochondrodysplasias (achondrogenesis type 1B, atelosteogenesis type II, and diastrophic dysplasia) with clinical features including but not limited to, dwarfism, spinal deformation, and specific joint abnormalities. Interestingly, while epidemiological studies have shown that the disease occurs in  
30 most populations, it is particularly prevalent in Finland owing to an apparent founder effect. For these reasons, siRNAs directed against this class of genes (including but not limited to SLC26A1-9, and related molecules) may be potentially helpful in both therapeutic and research venues.

## Ion Exchangers

Intracellular pH regulatory mechanisms are critical for the maintenance of countless cellular processes. For instance, in muscle cells, contractile processes and metabolic reactions are influenced by pH. During periods of increased energy 5 demands and ischemia, muscle cells produce large amounts of lactic acid that, without quick and efficient disposal, would lead to acidification of the sarcoplasm.

Several different transport mechanisms have evolved to maintain a relatively constant intracellular pH. The relative contribution of each of these processes varies 10 with cell type, the metabolic requirements of the cell, and the local environmental conditions. Intracellular pH regulatory processes that have been characterized functionally include but are not limited to the  $\text{Na}^+/\text{H}^+$  exchange, the  $\text{Na}(\text{HCO}_3)_n$  cotransport, and the  $\text{Na}^+$ -dependent and -independent  $\text{Cl}^-$ /base exchangers. As bicarbonate and  $\text{CO}_2$  comprise the major pH buffer of biological fluids, sodium 15 bicarbonate cotransporters (NBCs) are critical. Studies have shown that these molecules exist in numerous tissues including the kidney, brain, liver, cornea, heart, and lung, suggesting that NBCs play an important role in mediating  $\text{HCO}_3^-$  transport in both epithelial as well as nonepithelial cells.

Recent molecular cloning experiments have identified the existence of four NBC isoforms (NBC1, 2, 3 and 4) and two NBC-related proteins, AE4 and NCBE (Anion Exchanger 4 and Na-dependent Chloride-Bicarbonate Exchanger). The 20 secondary structure analyses and hydropathy profile of this family predict them to be intrinsic membrane proteins with 12 putative transmembrane domains and several family members exhibit *N*-linked glycosylation sites, protein kinases A and C, casein 25 kinase II, and ATP/GTP-binding consensus phosphorylation sites, as well as potential sites for myristylation and amidation. AE4 is a relatively recent addition to this family of proteins and shows between 30-48% homology with the other family members. When expressed in COS-7 cells and *Xenopus* oocytes AE4 exhibits sodium- 30 independent and DIDS-insensitive anion exchanger activity. Exchangers have been shown to be responsible for a variety of human diseases. For instance, mutations in three genes of the anion transporter family (SLC) are believed to cause known hereditary diseases, including chondrodysplasia (SLC26A2, DTD), diarrhea (A3, down-regulated in adenoma/chloride-losing diarrhea protein: DRA/CLD), and

goiter/deafness syndrome (A4, pendrin). Moreover, mutations in Na<sup>+</sup>/HCO<sub>3</sub> co-transporters have also been associated with various human maladies. For these reasons, siRNAs directed against these sorts of genes (e.g. SLC4A4-10, and related genes) may be useful for therapeutic and research purposes.

5

### Receptors Involved in Synaptic Transmission

In all vertebrates, fast inhibitory synaptic transmission is the result of the interaction between the neurotransmitters glycine (Gly) and  $\gamma$ -aminobutyric acid (GABA) and their respective receptors. The strychnine-sensitive glycine receptor is especially important in that it acts in the mammalian spinal cord and brain stem and has a well-established role in the regulation of locomotor behavior.

Glycine receptors display significant sequence homology to several other receptors including the nicotinic acetylcholine receptor, the aminobutyric acid receptor type A (GABA<sub>AR</sub>), and the serotonin receptor type 3 (5-HT<sub>3</sub>R) subunits. As members of the superfamily of ligand-gated ion channels, these polypeptides share common topological features. The glycine receptor is composed of two types of glycosylated integral membrane proteins ( $\alpha 1$ - $\alpha 4$  and  $\beta$ ) arranged in a pentameric suprastructure. The alpha subunit encodes a large extracellular, N-terminal domain that carries the structural determinants essential for agonist and antagonist binding, followed by four transmembrane spanning regions (TM1-TM4), with TM2 playing the critical role of forming the inner wall of the chloride channel.

The density, location, and subunit composition of glycine neurotransmitter receptors changes over the course of development. It has been observed that the amount of GlyR gene translation (assessed by the injection of developing rat cerebral cortex mRNA into *Xenopus* oocytes) decreases with age, whereas that of GABARs increases. In addition, the type and location of mRNAs coding for GlyR changes over the course of development. For instance in a study of the expression of alpha 1 and alpha 2 subunits in the rat, it was observed that (in embryonic periods E11-18) the mantle zone was scarce in the alpha 1 mRNA, but the germinal zone (matrix layer) at E11-14 expressed higher levels of the message. At postnatal day 0 (P0), the alpha 1 signals became manifested throughout the gray matter of the spinal cord. By contrast,

the spinal tissues at P0 exhibited the highest levels of alpha 2 mRNA, which decreased with the postnatal development.

In both, man and mouse mutant lines, mutations of GlyR subunit genes result  
5 in hereditary motor disorders characterized by exaggerated startle responses and increased muscle tone. Pathological alleles of the *Grla1* gene are associated with the murine phenotypes *oscillator* (*spd*<sup>ot</sup>) and *spasmodic* (*spd*). Similarly, a mutant allele of *Glrb* has been found to underly the molecular pathology of the *spastic* mouse (*spa*). Resembling the situation in the mouse, a variety of *GLRA1* mutant alleles have been  
10 shown to be associated with the human neurological disorder hyperekplexia or startle disease. For these reasons, siRNA directed against glycine receptors (GLRA1-3, GLRB, and related molecules), glutamate receptors, GABA receptors, ATP receptors, and related neurotransmitter receptor molecules may be valuable therapeutic and research reagents.

15

## Proteases

### Kallikreins

One important class of proteases are the kallikreins, serine endopeptidases that  
20 split peptide substrates preferentially on the C-terminal side of internal arginyl and lysyl residues. Kallikreins are generally divided into two distinct groups, plasma kallikreins and tissue kallikreins. Tissue kallikreins represent a large group of enzymes that have substantial similarities at both the gene and protein level. The genes encoding this group are frequently found on a single chromosome, are  
25 organized in clusters, and are expressed in a broad range of tissues (e.g. pancreas, ovaries, breast). In contrast, the plasma form of the enzyme is encoded by a single gene (e.g. KLK3) that has been localized to chromosome 4q34-35 in humans. The gene encoding plasma kallikrein is expressed solely in the liver, contains 15 exons, and encodes a glycoprotein that is translated as a preprotein called prekallikrein.

30

Kallikreins are believed to play an important role in a host of physiological events. For instance, the immediate consequence of plasma prekallikrein activation is the cleavage of high molecular weight kininogen (HK) and the subsequent liberation of bradykinin, a nine amino acid vasoactive peptide that is an important mediator of

inflammatory responses. Similarly, plasma kallikrein promotes single-chain urokinase activation and subsequent plasminogen activation, events that are critical to blood coagulation and wound healing.

5 Disruptions in the function of kallikreins have been implicated in a variety of pathological processes including imbalances in renal function and inflammatory processes. For these reasons, siRNAs directed against this class of genes (e.g. KLK1-15) may prove valuable in both research and therapeutic settings.

10 ADAM Proteins

The process of fertilization takes place in a series of discrete steps whereby the sperm interacts with; i) the cumulus cells and the hyaluronic acid extracellular matrix (ECM) in which they are embedded, ii) the egg's own ECM, called the *zona pellucida* (ZP), and iii) the egg plasma membrane. During the course of these interactions, the 15 "acrosome reaction," the exocytosis of the acrosome vesicle on the head of the sperm, is induced, allowing the sperm to penetrate the ZP and gain access to the perivitelline space. This process exposes new portions of the sperm membrane, including the inner acrosomal membrane and the equatorial segment, regions of the sperm head that can participate in initial gamete membrane binding.

20

The interactions of the gamete plasma membranes appear to involve multiple ligands and receptors and are frequently compared to leukocyte-endothelial interactions. These interactions lead to a series of signal transduction events in the egg, known as collectively as egg activation and include the initiation of oscillations 25 in intracellular calcium concentration, the exit from meiosis, the entry into the first embryonic mitosis, and the formation of a block to polyspermy via the release of ZP-modifying enzymes from the egg's cortical granules. Ultimately, sperm and egg not only adhere to each other but also go on to undergo membrane fusion, making one cell (the zygote) from two.

30

Studies on the process of sperm-egg interactions have identified a number of proteins that are crucial for fertilization. One class of proteins, called the ADAM family (A Disintegrin And Metalloprotease), has been found to be important in spermatogenesis and fertilization, as well as various developmental systems including

myogenesis and neurogenesis. Members of the family contain a disintegrin and metalloprotease domain (and therefore have (potentially) both cell adhesion and protease activities), as well as cysteine-rich regions, epidermal growth factor (EGF)-like domains, a transmembrane region, and a cytoplasmic tail. Currently, the ADAM 5 gene family has 29 members and constituents are widely distributed in many tissues including the brain, testis, epididymis, ovary, breast, placenta, liver, heart, lung, bone, and muscle.

One of the best-studied members of the ADAM family is fertilin, a 10 heterodimeric protein comprised of at least two subunits, fertilin alpha and fertilin beta. The fertilin beta gene (ADAM2) has been disrupted with a targeting gene construct corresponding to the exon encoding the fertilin beta disintegrin domain. Sperm from males homozygous for disruptions in this region exhibit defects in 15 multiple facets of sperm function including reduced levels of sperm transit from the uterus to the oviduct, reduced sperm-ZP binding, and reduced sperm-egg binding, all of which contribute to male infertility.

Recently, four new ADAM family members (ADAM 24-27) have been 20 isolated. The deduced amino acid sequences show that all four contain the complete domain organization common to ADAM family members and Northern Blot analysis has shown all four to be specific to the testes. SiRNAs directed against this class of genes (e.g. ADAM2 and related proteins) may be useful as research tools and therapeutics directed toward fertility and birth control.

25 Aminopeptidases

Aminopeptidases are proteases that play critical roles in processes such as protein maturation, protein digestion in its terminal stage, regulation of hormone levels, selective or homeostatic protein turnover, and plasmid stabilization. These enzymes generally have broad substrate specificity, occur in several forms and play a 30 major role in physiological homeostasis. For instance, the effects of bradykinin, angiotensin converting enzyme (ACE), and other vasoactive molecules are muted by one of several peptidases that cleave the molecule at an internal position and eliminate its ability to bind its cognate receptor (e.g. for bradykinin, the B2-receptor).

Among the enzymes that can cleave bradykinin is the membrane bound aminopeptidase P, also referred to as aminoacylproline aminopeptidase, proline aminopeptidase; X-Pro aminopeptidase (eukaryote) and XPNPEP2. Aminopeptidase P is an aminoacylproline aminopeptidase specific for NH<sub>2</sub>-terminal Xaa-proline bonds. The enzyme i) is a mono-zinc-containing molecule that lacks any of the typical metal binding motifs found in other zinc metalloproteases, ii) has an active-site configuration similar to that of other members of the MG peptidase family, and iii) is present in a variety of tissues including but not limited to the lung, kidney, brain, and intestine.

10

Aminopeptidases play an important role in a diverse set of human diseases. Low plasma concentrations of aminopeptidase P are a potential predisposing factor for development of angio-oedema in patients treated with ACE inhibitors, and inhibitors of aminopeptidase P may act as cardioprotectors against other forms of illness including, but not limited to myocardial infarction. For these reasons, siRNAs directed against this family of proteins (including but not limited to XPNPEP1 and related proteins) may be useful as research and therapeutic tools.

#### Serine Proteases

One important class of proteases are the serine proteases. Serine proteases share a common catalytic triad of three amino acids in their active site (serine (nucleophile), aspartate (electrophile), and histidine (base)) and can hydrolyze either esters or peptide bonds utilizing mechanisms of covalent catalysis and preferential binding of the transition state. Based on the position of their introns serine proteases have been classified into a minimum of four groups including those in which 1) the gene has no introns interrupting the exon coding for the catalytic triad (*e.g.* the haptoglobin gene); 2) each gene contains an intron just downstream from the codon for the histidine residue at the active site, a second intron downstream from the exon containing the aspartic acid residue of the active site and a third intron just upstream from the exon containing the serine of the active site (*e.g.* trypsinogen, chymotrypsinogen, kallikrein and proelastase); 3) the genes contain seven introns interrupting the exons coding the catalytic region (*e.g.* complement factor B gene); and 4) the genes contain two introns resulting in a large exon that contains both the

active site aspartic acid and serine residues (e.g. factor X, factor IX and protein C genes).

Cytotoxic lymphocytes (e.g. CD8(+) cytotoxic T cells and natural killer cells) 5 form the major defense of higher organisms against virus-infected and transformed cells. A key function of these cells is to detect and eliminate potentially harmful cells by inducing them to undergo apoptosis. This is achieved through two principal pathways, both of which require direct but transient contact between the killer cell and its target. The first pathway involves ligation of TNF receptor-like molecules such as 10 Fas/CD95 to their cognate ligands, and results in mobilization of conventional, programmed cell-death pathways centered on activation of pro-apoptotic caspases. The second mechanism consists of a pathway whereby the toxic contents of a specialized class of secretory vesicles are introduced into the target cell. Studies over the last two decades have identified the toxic components as Granzymes, a family of 15 serine proteases that are expressed exclusively by cytotoxic T lymphocytes and natural killer (NK) cells. These agents are stored in specialized lytic granules and enter the target cell via endocytosis. Like caspases, cysteine proteases that play an important role in apoptosis, granzymes can cleave proteins after acidic residues, especially aspartic acid, and induce apoptosis in the recipient cell.

20

Granzymes have been grouped into three subfamilies according to substrate specificity. Members of the granzyme family that have enzymatic activity similar to the serine protease chymotrypsin are encoded by a gene cluster termed the 'chymase locus'. Similarly, granzymes with trypsin-like specificities are encoded by the 25 'tryptase locus', and a third subfamily cleaves after unbranched hydrophobic residues, especially methionine, and are encoded by the 'Met-ase locus'. All granzymes are synthesized as zymogens and, after clipping of the leader peptide, obtain maximal enzymatic activity subsequent to the removal of an amino-terminal dipeptide.

30

Granzymes have been found to be important in a number of important biological functions including defense against intracellular pathogens, graft versus host reactions, the susceptibility to transplantable and spontaneous malignancies, lymphoid homeostasis, and the tendency toward auto-immune diseases. For these

reasons, siRNAs directed against granzymes (e.g. GZMA, GZMB, GZMH, GZHK, GZMM) and related serine proteases may be useful research and therapeutic reagents.

### Kinases

5           Protein Kinases (PKs) have been implicated in a number of biological processes. Kinase molecules play a central role in modulating cellular physiology and developmental decisions, and have been implicated in a large list of human maladies including cancer, diabetes, and others.

10           During the course of the last three decades, over a hundred distinct protein kinases have been identified, all with presumed specific cellular functions. A few of these enzymes have been isolated to sufficient purity to perform *in vitro* studies, but most remain intractable due to the low abundance of these molecules in the cell. To counter this technical difficulty, a number of protein kinases have been isolated by

15           molecular cloning strategies that utilize the conserved sequences of the catalytic domain to isolate closely related homologs. Alternatively, some kinases have been purified (and subsequently studied) based on their interactions with other molecules.

20           p58 is a member of the p34cdc2-related supergene family and contains a large domain that is highly homologous to the cell division control kinase, cdc2. This new cell division control-related protein kinase was originally identified as a component of semipurified galactosyltransferase; thus, it has been denoted galactosyltransferase-associated protein kinase (GTA-kinase). GTA-kinase has been found to be expressed in both adult and embryonic tissues and is known to phosphorylate a number of

25           substrates, including histone H1, and casein. Interestingly enough, over expression of this molecule in CHO cells has shown that elevated levels of p58 result in a prolonged late telophase and an early G1 phase, thus hinting of an important role for GTA-kinase in cell cycle regulation.

30           

### Cyclin Dependent Kinases

The cyclin-dependent kinases (Cdks) are a family of highly conserved serine/threonine kinases that mediate many of the cell cycle transitions that occur during duplication. Each of these Cdk catalytic subunits associates with a specific

subset of regulatory subunits, termed cyclins, to produce a distinct Cdk·cyclin kinase complex that, in general, functions to execute a unique cell cycle event.

Activation of the Cdk·cyclin kinases during cellular transitions is controlled by 5 a variety of regulatory mechanisms. For the Cdc2·cyclin B complex, inhibition of kinase activity during S phase and G<sub>2</sub> is accomplished by phosphorylation of two Cdc2 residues, Thr<sup>14</sup> and Tyr<sup>15</sup>, which are positioned within the ATP-binding cleft. Phosphorylation of Thr<sup>14</sup> and/or Tyr<sup>15</sup> suppresses the catalytic activity of the molecule by disrupting the orientation of the ATP present within this cleft. In contrast, the 10 abrupt dephosphorylation of these residues by the Cdc25 phosphatase results in the rapid activation of Cdc2·cyclin B kinase activity and subsequent downstream mitotic events. While the exact details of this pathway have yet to be elucidated, it has been proposed that Thr<sup>14</sup>/Tyr<sup>15</sup> phosphorylation functions to permit a cell to attain a critical concentration of inactive Cdk·cyclin complexes, which, upon activation, induces a 15 rapid and complete cell cycle transition. Furthermore, there is evidence in mammalian cells that Thr<sup>14</sup>/Tyr<sup>15</sup> phosphorylation also functions to delay Cdk activation after DNA damage.

The *Schizosaccharomyces pombe wee1* gene product was the first kinase 20 identified that is capable of phosphorylating Tyr<sup>15</sup> in Cdc2. Homologs of the Wee1 kinase have been subsequently identified and biochemically characterized from a wide range of species including human, mouse, frog, *Saccharomyces cerevisiae*, and *Drosophila*. In vertebrate systems, where Thr<sup>14</sup> in Cdc2 is also phosphorylated, the Wee1 kinase was capable of phosphorylating Cdc2 on Tyr<sup>15</sup>, but not Thr<sup>14</sup>, indicating 25 that another kinase was responsible for Thr<sup>14</sup> phosphorylation. This gene, Myt1 kinase, was recently isolated from the membrane fractions of *Xenopus* egg extracts and has been shown to be capable of phosphorylating Thr<sup>14</sup> and, to a lesser extent, Tyr<sup>15</sup> in Cdc2. A human Myt1 homolog displaying similar properties has been isolated, as well as a non-membrane-associated molecule with Thr<sup>14</sup> kinase activity.

30

In the past decade it has been shown that cancer can originate from overexpression of positive regulators, such as cyclins, or from underexpression of negative regulators (e.g. p16 (INK4a), p15 (INK4b), p21 (Cip1)). Inhibitors such as Myt1 are the focus of much cancer research because they are capable of controlling

cell cycle proliferation, now considered the Holy Grail for cancer treatment. For these reasons, siRNA directed against kinases and kinase inhibitors including but not limited to ABL1, ABL2, ACK1, ALK, AXL, BLK, BMX, BTK, C20orf64, CSF1R, SCK, DDR1, DDR2, DKFZp761P1010, EGFR, EPHA1, EPHA2, EPHA3, EPHA4, 5 EPHA7, EPHA8, EPHB1, EPHB2, EPHB3, EPHB4, EPHB6, ERBB2, ERBB3, ERBB4, FER, FES, FGFR1, FGFR2, FGFR3, FGFR4, FGR, FLT1, FLT3, FLT4, FRK, FYN, HCK, IGF1R, INSR, ITK, JAK1, JAK2, JAK3, KDR, KIAA1079, KIT, LCK, LTK, LYN, MATK, MERTK, MET, MST1R, MUSK, NTRK1, NTRK2, 10 NTRK3, PDGFRA, PDGFRB, PTK2, PTK2B, PTK6, PTK7, PTK9, PTK9L, RET, ROR1, ROR2, ROS1, RYK, SRC, SYK, TEC, TEK, TIE, TNK1, TXK, TYK2, TYRO3, YES1, and related proteins, may be useful for research and therapeutic purposes.

#### G Protein Coupled Receptors

15 One important class of genes to which siRNAs can be directed are G-protein coupled receptors (GPCRs). GPCRs constitute a superfamily of seven transmembrane spanning proteins that respond to a diverse array of sensory and chemical stimuli, such as light, odor, taste, pheromones, hormones and neurotransmitters. GPCRs play a 20 central role in cell proliferation, differentiation, and have been implicated in the etiology of disease.

The mechanism by which G protein-coupled receptors translate extracellular signals into cellular changes was initially envisioned as a simple linear model: 25 activation of the receptor by agonist binding leads to dissociation of the heterotrimeric GTP-binding G protein (Gs, Gi, or Gq) into its alpha and beta/gamma subunits, both of which can activate or inhibit various downstream effector molecules. More specifically, activation of the GPCR induces a conformational change in the G $\alpha$  subunit, causing GDP to be released and GTP to be bound in its place. The G $\alpha$  and 30 G $\beta\gamma$  subunits then dissociate from the receptor and interact with a variety of effector molecules. For instance in the case of the Gs family, the primary function is to stimulate the intracellular messenger adenylate cyclase (AC), which catalyzes the conversion of cytoplasmic ATP into the secondary messenger cyclic AMP (cAMP). In contrast, the Gi family inhibits this pathway and the Gq family activates

phospholipases C (PLC), which cleaves phosphatidylinositol 4,5, bisphosphate (PIP2) to generate inositol-1,4,5-phosphate (IP3) and diacylglycerol (DAG).

More recently, studies have shown that the functions of GPCRs are not limited to their actions on G-proteins and that considerable cross-talk exists between this diverse group of receptor molecules and a second class of membrane bound proteins, the receptor tyrosine kinases (RTKs). A number of GPCRs such as endothelin-1, thrombin, bombesin, and dopamine receptors can activate MAPKs, a downstream effector of the RTK/Ras pathway. Interestingly, the interaction between these two families is not unidirectional and RTKs can also modulate the activity of signaling pathways traditionally thought to be controlled exclusively by ligands that couple to GPCRs. For instance, EGF, which normally activates the MAPK cascade via the EGF receptor can stimulate adenylate cyclase activity by activating Gαs.

There are dozens of members of the G Protein-Coupled Receptor family that have emerged as prominent drug targets in the last decade. One non-limiting list of potential GPCR-siRNA targets is as follows:

#### CMKLR1

CML1/ CMKLR1 (Accession No. Q99788) is a member of the chemokine receptor family of GPCRs that may play a role in a number of diseases including those involved in inflammation and immunological responses (*e.g.* asthma, arthritis). For this reason, siRNA directed against this protein may prove to be important therapeutic reagents.

25

Studies of juvenile-onset neuronal ceroid lipofuscinosis (JNCL, Batten disease), the most common form of childhood encephalopathy that is characterized by progressive neural degeneration, show that it is brought on by mutations in a novel lysosomal membrane protein (CLN3). In addition to being implicated in JNCL, CLN3 (GPCR-like protein, Accession No. A57219) expression studies have shown that the CLN3 mRNA and protein are highly over-expressed in a number of cancers (*e.g.* glioblastomas, neuroblastomas, as well as cancers of the prostate, ovaries, breast, and colon) suggesting a possible contribution of this gene to tumor growth. For this

reason, siRNA directed against this protein may prove to be important therapeutic reagents.

### CLACR

5 The calcitonin receptor (CTR/ CALCR, Accession No. NM\_001742) belongs to "family B" of GPCRs which typically recognized regulatory peptides such as parathyroid hormone, secretin, glucagons and vasoactive intestinal polypeptide. Although the CT receptor typically binds to calcitonin (CT), a 32 amino acid peptide hormone produced primarily by the thyroid, association of the receptor with RAMP 10 (Receptor Activity Modulating Protein) enables it to readily bind other members of the calcitonin peptide family including amylin (AMY) and other CT gene-related peptides (e.g.  $\alpha$ CGRP and  $\beta$ CGRP). While the primary function of the calcitonin receptor pertains to regulating osteoclast mediated bone resorption and enhanced  $\text{Ca}^{+2}$  excretion by the kidney, recent studies have shown that CT and CTRs may play an 15 important role in a variety of processes as wide ranging as embryonic/foetal development and sperm function/physiology. In addition, studies have shown that patients with particular CTR genotypes may be at higher risk to lose bone mass and that this GPCR may contribute to the formation of calcium oxalate urinary stones. For this reason, siRNA directed against CTR may be useful as therapeutic reagents.

20

### OXTR

The human oxytocin receptor (OTR, OXTR) is a 389 amino acid polypeptide that exhibits the seven transmembrane domain structure and belongs to the Class-I (rhodopsin-type) family of G-protein coupled receptors. OTR is expressed in a wide 25 variety of tissues throughout development and mediates physiological changes through G(q) proteins and phospholipase C-beta. Studies on the functions of oxytocin and the oxytocin receptor have revealed a broad list of duties. OT and OTR play a role in a host of sexual, maternal and social behaviors that include egg-laying, birth, milk-letdown, feeding, grooming, memory and learning. In addition, it has been 30 hypothesized that abnormalities in the functionality of oxytocin-OTR receptor-ligand system can lead to a host of irregularities including compulsive behavior, eating disorders (such as anorexia), depression, and various forms of neurodegenerative

diseases. For these reasons, siRNA directed against this gene (NM\_000916) may play an important role in combating OTR-associated illnesses.

### EDG GPCRs

5 Lysophosphatidic acid and other lipid-based hormones/growth factors induce their effects by activating signaling pathways through the G-protein coupled receptors (GPCRs) and have been observed to play important roles in a number of human diseases including cancer, asthma, and vascular pathologies. For instance, during studies of immunoglobulin A nephropathy (IgAN), researchers have observed an  
10 enhanced expression of EDG5 (NP\_004221) suggesting a contribution of this gene product in the development of IgAN. For that reasons, siRNA directed against Edg5 (NM\_004230), Edg4 (NM\_004720), Edg7 (Nm\_012152) and related genes may play an important role in combating human disease.

### 15 Genes Involved in Cholesterol Signaling and Biosynthesis

Studies on model genetic organisms such as *Drosophila* and *C. elegans* have led to the identification of a plethora of genes that are essential for early development. Mutational analysis and ectopic expression studies have allowed many of these genes to be grouped into discreet signal transduction pathways and have shown that these  
20 elements play critical roles in pattern formation and cell differentiation. Disruption of one or more of these genes during early stages of development frequently leads to birth defects whereas as alteration of gene function at later stages in life can result in tumorigenesis.

25 One critical set of interactions known to exist in both invertebrates and vertebrates is the Sonic Hedgehog-Patched-Gli pathway. Originally documented as a *Drosophila* segmentation mutant, several labs have recently identified human and mouse orthologs of many of the pathways members and have successfully related disruptions in these genes to known diseases. Pathway activation is initiated with the  
30 secretion of Sonic hedgehog. There are three closely related members of the Shh family (Sonic hedgehog, Desert, and Indian) with Shh being the most widely expressed form of the group. The Shh gene product is secreted as a small pro-signal molecule. To successfully initiate its developmental role, Shh is first cleaved, whereupon the N-terminal truncated fragment is covalently modified with cholesterol.

The addition of the sterol moiety promotes the interaction between Shh and its cognate membrane bound receptor, Patched (Ptch). There are at least two isoforms of the Patched gene, Ptch1 and Ptch2. Both isoforms contain a sterol-sensing domain (SSD); a roughly 180 amino acid cluster that is found in at least seven different

5 classes of molecules including those involved in cholesterol biosynthesis, vesicular traffic, signal transduction, cholesterol transport, and sterol homeostasis. In the absence of Shh, the Patched protein is a negative regulator of the pathway. In contrast, binding of Shh-cholesterol to the Patched receptor releases the negative inhibition which that molecule enforces on a G-protein coupled receptor known as Smoothened.

10 Subsequent activation of Smoothened (directly or indirectly) leads to the triggering of a trio of transcription factors that belong to the Gli family. All three factors are relatively large, contain a characteristic C2-H2 zinc-finger pentamer, and recognize one of two consensus sequences (SEQ. ID NO. 0463 GACCACCCA or SEQ. ID NO. 0464 GAACCACCCA). In the absence of Shh, Gli proteins are cleaved by the

15 proteosome and the C-terminally truncated fragment translocates to the nucleus and acts as a dominant transcription repressor. In the presence of Shh-cholesterol, Gli repressor formation is inhibited and full-length Gli functions as a transcriptional activator.

20 Shh and other members of the Shh-PTCH-Gli pathway are expressed in a broad range of tissues (e.g. the notochord, the floorplate of the neural tube, the brain, and the gut) at early stages in development. Not surprisingly, mutations that lead to altered protein expression or function have been shown to induce developmental abnormalities. Defects in the human Shh gene have been shown to cause

25 holoprosencephaly, a midline defect that manifests itself as cleft lip or palate, CNS septation, and a wide range of other phenotypes. Interestingly, defects in cholesterol biosynthesis generate similar Shh-like disorders (e.g. Smith-Lemli-Opitz syndrome) suggesting that cholesterol modification of the Shh gene product is crucial for pathway function. Both the Patched and Smoothened genes have also been shown to

30 be clinically relevant with Smoothened now being recognized as an oncogene that, like PTCH-1 and PTCH-2, is believed to be the causative agent of several forms of adult tumors. For these reasons, siRNA directed against Smoothened (SMO, NM\_005631), Patched (PTCH, nm\_000264), and additional genes that participate in

cholesterol signaling, biosynthesis, and degradation, have potentially useful research and therapeutic applications.

Targeted Pathways.

5 In addition to targeting siRNA against one or more members of a family of proteins, siRNA can be directed against members of a pathway. Thus, for instance, siRNA can be directed against members of a signal transduction pathway (e.g. the insulin pathway, including AKT1-3, CBL, CBLB, EIF4EBP1, FOXO1A, FOXO3A, FRAP1, GSK3A, GSK3B, IGF1, IGF1R, INPP5D, INSR, IRS1, MLLT7, PDPK1, 10 PIK3CA, PIK3CB, PIK3R1, PIK3R2, PPP2R2B, PTEN, RPS6, RPS6KA1, RPX6KA3, SGK, TSC1, TSC2, AND XPO1), an apoptotic pathway (CASP3,6,7,8,9, DSH1/2, P110, P85, PDK1/2, CATEININ, HSP90, CDC37, P23, BAD, BCLXL, BCL2, SMAC, and others), pathways, involved in DNA damage, cell cycle, and other physiological (p53,MDM2, CHK1/2, BRCA1/2, ATM, ATR, P15INK4, P27, P21, 15 SKP2, CDC25C/A, 14-3-3, PLK, RB, CDK4, GLUT4, Inos, Mtor, FKBP, PPAR, RXR, ER). Similarly, genes involved in immune system function including TNFR1, IL-IR, IRAK1/2, TRAF2, TRAF6, TRADD, FADD, IKK $\epsilon$ , IKK $\gamma$ , IKK $\beta$ , IKK $\alpha$ , I $\kappa$ B $\alpha$ , I $\kappa$ B $\beta$ , p50, p65, Rac, RhoA, Cdc42, ROCK, Pak1/2/3/4/5/6, cIAP, HDAC1/2, CBP,  $\beta$ -TrCP, Rip2/4, and others are also important targets for the siRNAs described 20 in this document and may be useful in treating immune system disorders. Genes involved in apoptosis, such as Dsh1/2,PTEN, P110 (pan), P85, PDK1/2, Akt1, Akt2, Akt (pan), p70<sup>S6K</sup>, GSK3 $\beta$ , PP2A (cat),  $\beta$ -catenin, HSP90, Cdc37/p50, P23, Bad, BclxL, Bcl2, Smac/Diablo, and Ask1 are potentially useful in the treatment of diseases that involve defects in programmed cell death (e.g. cancer), while siRNA 25 agents directed against p53, MDM2, Chk1/2, BRCA1/2, ATM, ATR, p15<sup>INK4</sup>, P27, P21, Skp2, Cdc25C/A, 14-3-3 $\sigma/\epsilon$ , PLK, Rb, Cdk4, Glut4, iNOS, mTOR, FKBP, PPAR $\gamma$ , RXR $\alpha$ , ER $\alpha$  and related genes may play a critical role in combating diseases associated with disruptions in DNA repair, and cell cycle abnormalities.

30 Tables VI -Table X below provide examples of useful pools for inhibiting different genes in the human insulin pathway and tyrosine kinase pathways, proteins involved in the cell cycle, the production of nuclear receptors, and other genes. These particular pools are particularly useful in humans, but would be useful in any species

that generates an appropriately homologous mRNA. Further, within each of the listed pools any one sequence maybe used independently but preferably at least two of the listed sequences, more preferably at least three, and most preferably all of the listed sequences for a given gene is present.

5

Table VI

| Gene Name | Acc#      | GI       | L.L.  | Duplex #    | Sequence             | SEQ. ID NO |
|-----------|-----------|----------|-------|-------------|----------------------|------------|
| AKT1      | NM_005163 | 4885060  | 207   | D-003000-05 | GACAAGGACGGGCACATTA  | 465        |
| AKT1      | NM_005163 | 4885060  | 207   | D-003000-06 | GGACAAGGACGGGCACATT  | 466        |
| AKT1      | NM_005163 | 4885060  | 207   | D-003000-07 | GCTACTTCCTCCTCAAGAA  | 467        |
| AKT1      | NM_005163 | 4885060  | 207   | D-003000-08 | GACCGCCTCTGCTTTGTCA  | 468        |
| AKT2      | NM_001626 | 6715585  | 208   | D-003001-05 | GTACTTCGATGATGAATT   | 469        |
| AKT2      | NM_001626 | 6715585  | 208   | D-003001-06 | GCAAAGAGGGCATCAGTGA  | 470        |
| AKT2      | NM_001626 | 6715585  | 208   | D-003001-07 | GGGCTAAAGTGACCATGAA  | 471        |
| AKT2      | NM_001626 | 6715585  | 208   | D-003001-08 | GCAGAATGCCAGCTGATGA  | 472        |
| AKT3      | NM_005465 | 32307164 | 10000 | D-003002-05 | GGAGTAAACTGGCAAGATG  | 473        |
| AKT3      | NM_005465 | 32307164 | 10000 | D-003002-06 | GACATTAAATTTCCTCGAA  | 474        |
| AKT3      | NM_005465 | 32307164 | 10000 | D-003002-07 | GACCAAAGCCAACACATT   | 475        |
| CBL       |           |          |       | D-003002-08 | GAGGAGAGAATGAATTGTA  | 476        |
| CBL       | NM_005188 | 4885116  | 867   | D-003003-05 | GGAGACACATTCGGATT    | 477        |
| CBL       | NM_005188 | 4885116  | 867   | D-003003-06 | GATCTGACCTGCAATGATT  | 478        |
| CBL       | NM_005188 | 4885116  | 867   | D-003003-07 | GACAATCCCTCACAATAAA  | 479        |
| CBLB      |           |          |       | D-003003-08 | CCAGAAAGCTTGGTCATT   | 480        |
| CBLB      | NM_170662 | 29366807 | 868   | D-003004-05 | GACCATACCTCATAACAAG  | 481        |
| CBLB      | NM_170662 | 29366807 | 868   | D-003004-06 | TGAAAGACCTCCACCAATC  | 482        |
| CBLB      | NM_170662 | 29366807 | 868   | D-003004-07 | GATGAAGGCTCCAGGTGTT  | 483        |
| CBLB      | NM_170662 | 29366807 | 868   | D-003004-08 | TATCAGCATTACGACTTA   | 484        |
| EIF4EBP1  |           |          |       |             |                      |            |
| EIF4EBP1  | NM_004095 | 20070179 | 1978  | D-003005-05 | GCAATAGCCCAGAAGATAA  | 485        |
| EIF4EBP1  | NM_004095 | 20070179 | 1978  | D-003005-06 | CGCAATAGCCCAGAAGATA  | 486        |
| EIF4EBP1  | NM_004095 | 20070179 | 1978  | D-003005-07 | GAGATGGACATTAAAGCA   | 487        |
| EIF4EBP1  | NM_004095 | 20070179 | 1978  | D-003005-08 | CAATAGCCCAGAAGATAAG  | 488        |
| FOXO1A    |           |          |       |             |                      |            |
| FOXO1A    | NM_002015 | 9257221  | 2308  | D-003006-05 | CCAGGCATCTCATAACAAA  | 489        |
| FOXO1A    | NM_002015 | 9257221  | 2308  | D-003006-06 | CCAGATGCCTATACAAACA  | 490        |
| FOXO1A    | NM_002015 | 9257221  | 2308  | D-003006-07 | GGAGGTATGAGTCAGTATA  | 491        |
| FOXO1A    | NM_002015 | 9257221  | 2308  | D-003006-08 | GAGGTATGAGTCAGTATAA  | 492        |
| FOXO3A    |           |          |       |             |                      |            |
| FOXO3A    | NM_001455 | 4503738  | 2309  | D-003007-01 | CAATAGCAACAAGTATACC  | 493        |
| FOXO3A    | NM_001455 | 4503738  | 2309  | D-003007-02 | TGAAGTCCAGGACGATGAT  | 494        |
| FOXO3A    | NM_001455 | 4503738  | 2309  | D-003007-03 | TGTCACACTATGGTAACCA  | 495        |
| FOXO3A    | NM_001455 | 4503738  | 2309  | D-003007-04 | TGTTCAATGGGAGCTTGG   | 496        |
| FRAP1     |           |          |       |             |                      |            |
| FRAP1     | NM_004958 | 19924298 | 2475  | D-003008-05 | GAGAAGAAATGGAAGAAAT  | 497        |
| FRAP1     | NM_004958 | 19924298 | 2475  | D-003008-06 | CCAAAGTGCTGCAGTACTA  | 498        |
| FRAP1     | NM_004958 | 19924298 | 2475  | D-003008-07 | GAGCATGCCGTCAATAATA  | 499        |
| FRAP1     | NM_004958 | 19924298 | 2475  | D-003008-08 | GGTCTGAACGTGAATGAAGA | 500        |

|        |           |          |      |             |                      |     |
|--------|-----------|----------|------|-------------|----------------------|-----|
| GSK3A  |           |          |      |             |                      |     |
| GSK3A  | NM_019884 | 11995473 | 2931 | D-003009-05 | GGACAAAGGTGTTCAAATC  | 501 |
| GSK3A  | NM_019884 | 11995473 | 2931 | D-003009-06 | GAACCCAGCTGCTAACAA   | 502 |
| GSK3A  | NM_019884 | 11995473 | 2931 | D-003009-07 | GCGCACAGCTTCTTGATG   | 503 |
| GSK3A  | NM_019884 | 11995473 | 2931 | D-003009-08 | GCTCTAGCCTGCTGGAGTA  | 504 |
| GSK3B  |           |          |      |             |                      |     |
| GSK3B  | NM_002093 | 21361339 | 2932 | D-003010-05 | GAAGAAAGATGAGGTCTAT  | 505 |
| GSK3B  | NM_002093 | 21361339 | 2932 | D-003010-06 | GGACCCAAATGTCAAACCA  | 506 |
| GSK3B  | NM_002093 | 21361339 | 2932 | D-003010-07 | GAAATGAACCCAAACTACA  | 507 |
| GSK3B  | NM_002093 | 21361339 | 2932 | D-003010-08 | GATGAGGGTCTATCTTAATC | 508 |
| IGF1   |           |          |      |             |                      |     |
| IGF1   | NM_000618 |          |      | D-003011-05 | GGAAGTACATTGAAGAAC   | 509 |
| IGF1   | NM_000618 |          |      | D-003011-06 | AGAAGGAAGTACATTGAA   | 510 |
| IGF1   | NM_000618 |          |      | D-003011-07 | CCTCAAGCCTGCCAAGTCA  | 511 |
| IGF1   | NM_000618 |          |      | D-003011-08 | GGTGGATGCTCTTCAGTTC  | 512 |
| IGF1R  |           |          |      |             |                      |     |
| IGF1R  | NM_000875 | 11068002 | 3480 | D-003012-05 | CAACGAAGCTCTGTGATG   | 513 |
| IGF1R  | NM_000875 | 11068002 | 3480 | D-003012-06 | GGCCAGAAATGGAGAATAA  | 514 |
| IGF1R  | NM_000875 | 11068002 | 3480 | D-003012-07 | GAAGCACCCCTTAAGAATG  | 515 |
| IGF1R  | NM_000875 | 11068002 | 3480 | D-003012-08 | GCAGACACCTACAACATCA  | 516 |
| INPP5D |           |          |      |             |                      |     |
| INPP5D | NM_005541 | 5031798  | 3635 | D-003013-05 | GGAATTGCGTTTACACTTA  | 517 |
| INPP5D | NM_005541 | 5031798  | 3635 | D-003013-06 | GGAAACTGATCATTAAGAA  | 518 |
| INPP5D | NM_005541 | 5031798  | 3635 | D-003013-07 | CGACAGGGATGAAGTACAA  | 519 |
| INPP5D | NM_005541 | 5031798  | 3635 | D-003013-08 | AAACGCAGCTGCCATCTA   | 520 |
| INSR   |           |          |      |             |                      |     |
| INSR   | NM_000208 | 4557883  | 3643 | D-003014-05 | GGAAGACGTTGAGGATTA   | 521 |
| INSR   | NM_000208 | 4557883  | 3643 | D-003014-06 | GAACAAGGCTCCGAGAGT   | 522 |
| INSR   | NM_000208 | 4557883  | 3643 | D-003014-07 | GGAGAGACCTTGGAAATTG  | 523 |
| INSR   | NM_000208 | 4557883  | 3643 | D-003014-08 | GGACGGAACCCACCTATT   | 524 |
| IRS1   |           |          |      |             |                      |     |
| IRS1   | NM_005544 | 5031804  | 3667 | D-003015-05 | AAAGAGGTCTGGCAAGTGA  | 525 |
| IRS1   | NM_005544 | 5031804  | 3667 | D-003015-06 | GAACCTGATTGGTATCTAC  | 526 |
| IRS1   | NM_005544 | 5031804  | 3667 | D-003015-07 | CCACGGCGATCTAGTGCCT  | 527 |
| IRS1   | NM_005544 | 5031804  | 3667 | D-003015-08 | GTCAGTCTGTCGTCCAGTA  | 528 |
| MLLT7  |           |          |      |             |                      |     |
| MLLT7  | NM_005938 | 5174578  | 4303 | D-003016-05 | GGACTGGACTTCAACTTG   | 529 |
| MLLT7  | NM_005938 | 5174578  | 4303 | D-003016-06 | CCACGAAGCAGTTCAAATG  | 530 |
| MLLT7  | NM_005938 | 5174578  | 4303 | D-003016-07 | GAGAACGACTGACACTTG   | 531 |
| MLLT7  | NM_005938 | 5174578  | 4303 | D-003016-08 | GACCAGAGATCGCTAACCA  | 532 |
| PDPK1  |           |          |      |             |                      |     |
| PDPK1  | NM_002613 | 4505694  | 5170 | D-003017-05 | CAAGAGACCTCGTGGAGAA  | 533 |
| PDPK1  | NM_002613 | 4505694  | 5170 | D-003017-06 | GACCAGAGGCCAAGAATT   | 534 |
| PDPK1  | NM_002613 | 4505694  | 5170 | D-003017-07 | GGAAACGAGTATCTTATAT  | 535 |
| PDPK1  | NM_002613 | 4505694  | 5170 | D-003017-08 | GAGAACGACATATCATAA   | 536 |
| PIK3CA |           |          |      |             |                      |     |
| PIK3CA | NM_006218 | 5453891  | 5290 | D-003018-05 | GCTATCATCTGAACAATT   | 537 |
| PIK3CA | NM_006218 | 5453891  | 5290 | D-003018-06 | GGATAGAGGCCAATAATA   | 538 |
| PIK3CA | NM_006218 | 5453891  | 5290 | D-003018-07 | GGACAACTGTTCATATAG   | 539 |
| PIK3CA | NM_006218 | 5453891  | 5290 | D-003018-08 | GCCAGTACCTCATGGATTA  | 540 |
| PIK3CB |           |          |      |             |                      |     |
| PIK3CB | NM_006219 | 5453893  | 5291 | D-003019-05 | CGACAAAGACTGCCGAGAGA | 541 |
| PIK3CB | NM_006219 | 5453893  | 5291 | D-003019-06 | TCAAGTGTCTCTTAATATG  | 542 |
| PIK3CB | NM_006219 | 5453893  | 5291 | D-003019-07 | GGATTCAGTTGGAGTGATT  | 543 |
| PIK3CB | NM_006219 | 5453893  | 5291 | D-003019-08 | TTTCAAGTGTCTCCTAATA  | 544 |
| PIK3R1 |           |          |      |             |                      |     |

|         |           |          |      |             |                        |     |
|---------|-----------|----------|------|-------------|------------------------|-----|
| PIK3R1  | NM_181504 | 32455251 | 5295 | D-003020-05 | GGAAATATGGCTTCTCTGA    | 545 |
| PIK3R1  | NM_181504 | 32455251 | 5295 | D-003020-06 | GAAAGACGAGAGACCAATA    | 546 |
| PIK3R1  | NM_181504 | 32455251 | 5295 | D-003020-07 | GTAAAGCATTGTGTCAAA     | 547 |
| PIK3R1  | NM_181504 | 32455251 | 5295 | D-003020-08 | GGATCAAGTTGTCAAAGAA    | 548 |
| PIK3R2  |           |          |      |             |                        |     |
| PIK3R2  | NM_005027 | 4826907  | 5296 | D-003021-05 | GGAAAGGCGGGAACAAATAA   | 549 |
| PIK3R2  | NM_005027 | 4826907  | 5296 | D-003021-06 | GATGAAGCGTACTGCAATT    | 550 |
| PIK3R2  | NM_005027 | 4826907  | 5296 | D-003021-07 | GGACAGCGAACATCTCACTAC  | 551 |
| PIK3R2  | NM_005027 | 4826907  | 5296 | D-003021-08 | GCAAGAGATCCGAGACCAAGTA | 552 |
| PPP2R2B |           |          |      |             |                        |     |
| PPP2R2B | NM_004576 | 4758953  | 5521 | D-003022-05 | GAATGCAGCTTACTTCTT     | 553 |
| PPP2R2B | NM_004576 | 4758953  | 5521 | D-003022-06 | GACCGAAGCTGACATTATC    | 554 |
| PPP2R2B | NM_004576 | 4758953  | 5521 | D-003022-07 | TCGATTACCTGAAGAGTTT    | 555 |
| PPP2R2B | NM_004576 | 4758953  | 5521 | D-003022-08 | CCTGAAGAGTTAGAAATA     | 556 |
| PTEN    |           |          |      |             |                        |     |
| PTEN    | NM_000314 | 4506248  | 5728 | D-003023-05 | GTGAAGATCTTGACCAATG    | 557 |
| PTEN    | NM_000314 | 4506248  | 5728 | D-003023-06 | GATCAGCATAACACAAATTA   | 558 |
| PTEN    | NM_000314 | 4506248  | 5728 | D-003023-07 | GGCGCTATGTGTATTATTA    | 559 |
| PTEN    | NM_000314 | 4506248  | 5728 | D-003023-08 | GTATAGAGCGTGCAGATAA    | 560 |
| RPS6    |           |          |      |             |                        |     |
| RPS6    | NM_001010 | 17158043 | 6194 | D-003024-05 | GCCAGAAACTCATTGAAGT    | 561 |
| RPS6    | NM_001010 | 17158043 | 6194 | D-003024-06 | GGATATTCTGGACTGACT     | 562 |
| RPS6    | NM_001010 | 17158043 | 6194 | D-003024-07 | CCAAGGAGAACTGGAGAAA    | 563 |
| RPS6    | NM_001010 | 17158043 | 6194 | D-003024-08 | GCGTATGGCCACAGAAAGTT   | 564 |
| RPS6KA1 |           |          |      |             |                        |     |
| RPS6KA1 | NM_002953 | 20149546 | 6195 | D-003025-05 | GATGACACCTTCTACTTTG    | 565 |
| RPS6KA1 | NM_002953 | 20149546 | 6195 | D-003025-06 | GAGAATGGGCTCCTCATGA    | 566 |
| RPS6KA1 | NM_002953 | 20149546 | 6195 | D-003025-07 | CAAGCGGGATCCTTCAGAA    | 567 |
| RPS6KA1 | NM_002953 | 20149546 | 6195 | D-003025-08 | CCACCGGCCTGATGGAAGA    | 568 |
| RPS6KA3 |           |          |      |             |                        |     |
| RPS6KA3 | NM_004586 | 4759049  | 6197 | D-003026-05 | GAAGGGAAAGTTGTATCTTA   | 569 |
| RPS6KA3 | NM_004586 | 4759049  | 6197 | D-003026-06 | GAAAGTATGTGTATGTAGT    | 570 |
| RPS6KA3 | NM_004586 | 4759049  | 6197 | D-003026-07 | GGACAGCATCCAAACATTA    | 571 |
| RPS6KA3 | NM_004586 | 4759049  | 6197 | D-003026-08 | GGAGGTGAATTGCTGGATA    | 572 |
| SGK     |           |          |      |             |                        |     |
| SGK     | NM_005627 | 5032090  | 6446 | D-003027-01 | TTAATGGTGGAGAGTTGTT    | 573 |
| SGK     | NM_005627 | 5032090  | 6446 | D-003027-04 | ATTAACCTGGGATGATCTCA   | 574 |
| SGK     | NM_005627 | 25168262 | 6446 | D-003027-05 | GAAGAAAGCAATCCTGAAA    | 575 |
| SGK     | NM_005627 | 25168262 | 6446 | D-003027-06 | AAACACAGCTGAAATGTAC    | 576 |
| TSC1    |           |          |      |             |                        |     |
| TSC1    | NM_000368 | 24475626 | 7248 | D-003028-05 | GAAGATGGCTATTCTGTGT    | 577 |
| TSC1    | NM_000368 | 24475626 | 7248 | D-003028-06 | TATGAAGGCTCGAGAGTTA    | 578 |
| TSC1    | NM_000368 | 24475626 | 7248 | D-003028-07 | CGACACGGCTGATAACTGA    | 579 |
| TSC1    | NM_000368 | 24475626 | 7248 | D-003028-08 | CGGCTGATGTTGTTAAATA    | 580 |
| TSC2    |           |          |      |             |                        |     |
| TSC2    | NM_000548 | 10938006 | 7249 | D-003029-05 | GCATTAATCTCTTACCATTA   | 581 |
| TSC2    | NM_000548 | 10938006 | 7249 | D-003029-06 | CCAATGTCCTTGTCTTT      | 582 |
| TSC2    | NM_000548 | 10938006 | 7249 | D-003029-07 | GGAGACACATCACCTACTT    | 583 |
| TSC2    | NM_000548 | 10938006 | 7249 | D-003029-08 | TCACCAAGGCTCATCAAGAA   | 584 |
| XPO1    |           |          |      |             |                        |     |
| XPO1    | NM_003400 | 8051634  | 7514 | D-003030-05 | GAAAGTCTGTCAAATAA      | 585 |
| XPO1    | NM_003400 | 8051634  | 7514 | D-003030-06 | GCAATAGGCTCCATTAGTG    | 586 |
| XPO1    | NM_003400 | 8051634  | 7514 | D-003030-07 | GGAACATGATCAACTTATA    | 587 |
| XPO1    | NM_003400 | 8051634  | 7514 | D-003030-08 | GGATACAGATTCCATAAAAT   | 588 |

Table VII

| Gene Name | Acc#      | GI       | L.L   | Duplex #    | Sequence             | Se ID |
|-----------|-----------|----------|-------|-------------|----------------------|-------|
| ABL1      |           |          |       |             |                      |       |
| ABL1      | NM_007313 | 6382057  | 25    | D-003100-05 | GGAAATCAGTGACATAGTG  | 5     |
| ABL1      | NM_007313 | 6382057  | 25    | D-003100-06 | GGTCCACACTGCAATGTTT  | 5     |
| ABL1      | NM_007313 | 6382057  | 25    | D-003100-07 | GAAGGAAATCAGTGACATA  | 5     |
| ABL1      | NM_007313 | 6382057  | 25    | D-003100-08 | TCACTGAGTTCATGACCTA  | 5     |
| ABL2      |           |          |       |             |                      |       |
| ABL2      | NM_007314 | 6382061  | 27    | D-003101-05 | GAAATGGAGCGAACAGATA  | 5     |
| ABL2      | NM_007314 | 6382061  | 27    | D-003101-06 | GAGCCAAATTCCTATTAA   | 5     |
| ABL2      | NM_007314 | 6382061  | 27    | D-003101-07 | GTAATAAGCCTACAGTCTA  | 5     |
| ABL2      | NM_007314 | 6382061  | 27    | D-003101-08 | GGAGTGAAGTTCGCTCTAA  | 5     |
| ACK1      |           |          |       |             |                      |       |
| ACK1      | NM_005781 | 8922074  | 10188 | D-003102-05 | AAACGCAAGTCGTGGATGA  | 5     |
| ACK1      | NM_005781 | 8922074  | 10188 | D-003102-06 | GCAAGTCGTGGATGAGTAA  | 5     |
| ACK1      | NM_005781 | 8922074  | 10188 | D-003102-07 | GAGCACTACCTCAGAATGA  | 5     |
| ACK1      | NM_005781 | 8922074  | 10188 | D-003102-08 | TCAGCAGCACCCACTATTAA | 6     |
| ALK       |           |          |       |             |                      |       |
| ALK       | NM_004304 | 29029631 | 238   | D-003103-05 | GACAAGATCCTGCAGAATA  | 6     |
| ALK       | NM_004304 | 29029631 | 238   | D-003103-06 | GGAAGAGTCTGGCAGTTGA  | 6     |
| ALK       | NM_004304 | 29029631 | 238   | D-003103-07 | GCACGTGGCTCGGGACATT  | 6     |
| ALK       | NM_004304 | 29029631 | 238   | D-003103-08 | GAACTGCAGTGAAGGAACA  | 6     |
| AXL       |           |          |       |             |                      |       |
| AXL       | NM_021913 | 21536465 | 558   | D-003104-05 | GGTCAGAGCTGGAGGATT   | 6     |
| AXL       | NM_021913 | 21536465 | 558   | D-003104-06 | GAAAGAAGGAGACCCGTTA  | 6     |
| AXL       | NM_021913 | 21536465 | 558   | D-003104-07 | CCAAGAAGATCTACAATGG  | 6     |
| AXL       | NM_021913 | 21536465 | 558   | D-003104-08 | GGAACATGCATGCTGAATGA | 6     |
| BLK       |           |          |       |             |                      |       |
| BLK       | NM_001715 | 4502412  | 640   | D-003105-05 | GAGGATGCCTGCTGGATT   | 6     |
| BLK       | NM_001715 | 4502412  | 640   | D-003105-06 | ACATGAAGGTGGCCATTAA  | 6     |
| BLK       | NM_001715 | 4502412  | 640   | D-003105-07 | GGTCAGCGCCCAAGACAAG  | 6     |
| BLK       | NM_001715 | 4502412  | 640   | D-003105-08 | GAAACTCGGGCTGGACAA   | 6     |
| BMX       |           |          |       |             |                      |       |
| BMX       | NM_001721 | 21359831 | 660   | D-003106-05 | AAACAAACCTTCCTACTA   | 6     |
| BMX       | NM_001721 | 21359831 | 660   | D-003106-06 | GAAGGAGCATTATGGTTA   | 6     |
| BMX       | NM_001721 | 21359831 | 660   | D-003106-07 | GAGAAGAGATTACCTTGTT  | 6     |
| BMX       | NM_001721 | 21359831 | 660   | D-003106-08 | GTAAGGCTGTGAATGATAA  | 6     |
| BTK       |           |          |       |             |                      |       |
| BTK       | NM_000061 | 4557376  | 695   | D-003107-05 | GAACAGGAATGGAAGCTTA  | 6     |
| BTK       | NM_000061 | 4557376  | 695   | D-003107-06 | GCTATGGGCTGCCAAATT   | 6     |
| BTK       | NM_000061 | 4557376  | 695   | D-003107-07 | GAAAGCAACTTACCATGGT  | 6     |
| BTK       | NM_000061 | 4557376  | 695   | D-003107-08 | GGTAAACGATCAAGGAGTT  | 6     |
| C20orf64  |           |          |       |             |                      |       |
| C20orf64  | NM_033550 | 19923655 | 11285 | D-003108-05 | CAACTTAGCCAAGACAATT  | 6     |
| C20orf64  | NM_033550 | 19923655 | 11285 | D-003108-06 | GAAATTGAAGGCTCAGTGA  | 6     |
| C20orf64  | NM_033550 | 19923655 | 11285 | D-003108-07 | TGGAACAGCTGAACATTGT  | 6     |
| C20orf64  | NM_033550 | 19923655 | 11285 | D-003108-08 | GCTTCCAAC TGCTTATATA | 6     |
| CSF1R     |           |          |       |             |                      |       |
| CSF1R     | NM_005211 | 27262658 | 1436  | D-003109-05 | GGAGAGCTCTGACGTTGA   | 6     |
| CSF1R     | NM_005211 | 27262658 | 1436  | D-003109-06 | CAACAACGCTACCTTCCAA  | 6     |
| CSF1R     | NM_005211 | 27262658 | 1436  | D-003109-07 | CCACGCAGCTGCCTTACAA  | 6     |
| CSF1R     | NM_005211 | 27262658 | 1436  | D-003109-08 | GGAAACAACCTGCAGTTGG  | 6     |
| CSK       |           |          |       |             |                      |       |

|                   |           |          |       |             |                      |   |
|-------------------|-----------|----------|-------|-------------|----------------------|---|
| CSK               | NM_004383 | 4758077  | 1445  | D-003110-05 | CAGAATGTATTGCCAAGTA  | 6 |
| CSK               | NM_004383 | 4758077  | 1445  | D-003110-06 | GAACAAAGTCGCCGTCAAG  | 6 |
| CSK               | NM_004383 | 4758077  | 1445  | D-003110-07 | GCGAGTGCCTTATCCAAGA  | 6 |
| CSK               | NM_004383 | 4758077  | 1445  | D-003110-08 | GGAGAAGGGCTACAAGATG  | 6 |
| DDR1              |           |          |       |             |                      |   |
| DDR1              | NM_013994 | 7669484  | 780   | D-003111-05 | GGAGATGGAGTTGAGTTT   | 6 |
| DDR1              | NM_013994 | 7669484  | 780   | D-003111-06 | CAGAGGCCCTGTCATCTTT  | 6 |
| DDR1              | NM_013994 | 7669484  | 780   | D-003111-07 | GCTGGTAGCTGTCAAGATC  | 6 |
| DDR1              | NM_013994 | 7669484  | 780   | D-003111-08 | TGAAAGAGGTGAAGATCAT  | 6 |
| DDR2              |           |          |       |             |                      |   |
| DDR2              | NM_006182 | 5453813  | 4921  | D-003112-05 | GGTAAGAACTACACAATCA  | 6 |
| DDR2              | NM_006182 | 5453813  | 4921  | D-003112-06 | GAACGAGAGTGCCACCAAT  | 6 |
| DDR2              | NM_006182 | 5453813  | 4921  | D-003112-07 | ACACCAATCTGAAGTTAT   | 6 |
| DDR2              | NM_006182 | 5453813  | 4921  | D-003112-08 | CAACAAGAATGCCAGGAAT  | 6 |
| DKFZp761<br>P1010 |           |          |       |             |                      |   |
| DKFZp761<br>P1010 | NM_018423 | 8922178  | 55359 | D-003113-05 | CCTAGAAGCTGCCATTAAA  | 6 |
| DKFZp761<br>P1010 | NM_018423 | 8922178  | 55359 | D-003113-06 | GATTAGGCCTGGCTTATGA  | 6 |
| DKFZp761<br>P1010 | NM_018423 | 8922178  | 55359 | D-003113-07 | CCCAGTAGCTGCACACATA  | 6 |
| DKFZp761<br>P1010 | NM_018423 | 8922178  | 55359 | D-003113-08 | GGTGGTACCTGAACTGTAT  | 6 |
| EGFR              |           |          |       |             |                      |   |
| EGFR              | NM_005228 | 4885198  | 1956  | D-003114-05 | GAAGGAAACTGAATTCAAA  | 6 |
| EGFR              | NM_005228 | 4885198  | 1956  | D-003114-06 | GGAAATATGTAATCAGAAA  | 6 |
| EGFR              | NM_005228 | 4885198  | 1956  | D-003114-07 | CCACAAAGCAGTGAATTAA  | 6 |
| EGFR              | NM_005228 | 4885198  | 1956  | D-003114-08 | GTAACAAGCTCACGCAGTT  | 6 |
| EPHA1             |           |          |       |             |                      |   |
| EPHA1             | NM_005232 | 4885208  | 2041  | D-003115-05 | GACCAGAGCTTCACCATT   | 6 |
| EPHA1             | NM_005232 | 4885208  | 2041  | D-003115-06 | GCAAGACTGTGGCCATTAA  | 6 |
| EPHA1             | NM_005232 | 4885208  | 2041  | D-003115-07 | GGCGAACCTGACCTATGA   | 6 |
| EPHA1             | NM_005232 | 4885208  | 2041  | D-003115-08 | GATTGTAGCCGTATCTTT   | 6 |
| EPHA2             |           |          |       |             |                      |   |
| EPHA2             | NM_004431 | 4758277  | 1969  | D-003116-05 | GGAGGGATCTGGCAACTTG  | 6 |
| EPHA2             | NM_004431 | 4758277  | 1969  | D-003116-06 | GCAGCAAGGTGCACGAATT  | 6 |
| EPHA2             | NM_004431 | 4758277  | 1969  | D-003116-07 | GGAGAAGGATGGCGAGTTC  | 6 |
| EPHA2             | NM_004431 | 4758277  | 1969  | D-003116-08 | GAAGTTCACTACCGAGATC  | 6 |
| EPHA3             |           |          |       |             |                      |   |
| EPHA3             | NM_005233 | 21361240 | 2042  | D-003117-05 | GATCGGACCTCCAGAAATA  | 6 |
| EPHA3             | NM_005233 | 21361240 | 2042  | D-003117-06 | GAACTCAGCTCAGAAGATT  | 6 |
| EPHA3             | NM_005233 | 21361240 | 2042  | D-003117-07 | GCAAGAGGCACAAATGTTA  | 6 |
| EPHA3             | NM_005233 | 21361240 | 2042  | D-003117-08 | GAGCATCAGTTACAAAGA   | 6 |
| EPHA4             |           |          |       |             |                      |   |
| EPHA4             | NM_004438 | 4758279  | 2043  | D-003118-05 | GGTCTGGGATGAAGTATT   | 6 |
| EPHA4             | NM_004438 | 4758279  | 2043  | D-003118-06 | GAATGAAGTTACCTTATTG  | 6 |
| EPHA4             | NM_004438 | 4758279  | 2043  | D-003118-07 | GAACCTGGGTGGATAGCAA  | 6 |
| EPHA4             | NM_004438 | 4758279  | 2043  | D-003118-08 | GAGATTAAATTCACCTTGA  | 6 |
| EPHA7             |           |          |       |             |                      |   |
| EPHA7             | NM_004440 | 4758281  | 2045  | D-003119-05 | GAAAAGAGATGTTGCAGTA  | 6 |
| EPHA7             | NM_004440 | 4758281  | 2045  | D-003119-06 | CTAGATGCCCTCTGTATTAA | 6 |
| EPHA7             | NM_004440 | 4758281  | 2045  | D-003119-07 | AGAAGAAGGTTATCGTTA   | 6 |
| EPHA7             | NM_004440 | 4758281  | 2045  | D-003119-08 | TAGCAAAGCTGACCAAGAA  | 6 |
| EPHA8             |           |          |       |             |                      |   |
| EPHA8             | NM_020526 | 18201903 | 2046  | D-003120-05 | GAAGATGCACTATCAGAAT  | 6 |

|       |           |          |      |             |                      |   |
|-------|-----------|----------|------|-------------|----------------------|---|
| EPHA8 | NM_020526 | 18201903 | 2046 | D-003120-06 | GAGAAGATGCACTATCAGA  | € |
| EPHA8 | NM_020526 | 18201903 | 2046 | D-003120-07 | AACCTGATCTCCAGTGTGA  | € |
| EPHA8 | NM_020526 | 18201903 | 2046 | D-003120-08 | TCTCAGACCTGGGCTATGT  | € |
| EPHB1 |           |          |      |             |                      |   |
| EPHB1 | NM_004441 | 21396502 | 2047 | D-003121-05 | GCGATAAGCTCCAGCATT   | € |
| EPHB1 | NM_004441 | 21396502 | 2047 | D-003121-06 | GAAACGGGCTTATAGCAAA  | € |
| EPHB1 | NM_004441 | 21396502 | 2047 | D-003121-07 | GGATGAAGATCTACATTGA  | € |
| EPHB1 | NM_004441 | 21396502 | 2047 | D-003121-08 | GCACGTCTGTCAACATC    | € |
| EPHB2 |           |          |      |             |                      |   |
| EPHB2 | NM_017449 | 17975764 | 2048 | D-003122-05 | ACTATGAGCTGCAGTACTA  | € |
| EPHB2 | NM_017449 | 17975764 | 2048 | D-003122-06 | GTACAACGCCACAGCCATA  | € |
| EPHB2 | NM_017449 | 17975764 | 2048 | D-003122-07 | GGAAAGCAATGACTGTTCT  | € |
| EPHB2 | NM_017449 | 17975764 | 2048 | D-003122-08 | CGGACAAGCTGCAACACTA  | € |
| EPHB3 |           |          |      |             |                      |   |
| EPHB3 | NM_004443 | 17975767 | 2049 | D-003123-05 | GGTGTGATCTCCAATGTGA  | € |
| EPHB3 | NM_004443 | 17975767 | 2049 | D-003123-06 | GGGATGACCTCCTGTACAA  | € |
| EPHB3 | NM_004443 | 17975767 | 2049 | D-003123-07 | CAGAAGACCTGCTCCGTAT  | € |
| EPHB3 | NM_004443 | 17975767 | 2049 | D-003123-08 | GAGATGAAGTACTTTGAGA  | € |
| EPHB4 |           |          |      |             |                      |   |
| EPHB4 | NM_004444 | 17975769 | 2050 | D-003124-05 | GGACAAACACGGACAGTAT  | € |
| EPHB4 | NM_004444 | 17975769 | 2050 | D-003124-06 | GTACTAAGGTCTACATCGA  | € |
| EPHB4 | NM_004444 | 17975769 | 2050 | D-003124-07 | GGAGAGAAGCAGAATATT   | € |
| EPHB4 | NM_004444 | 17975769 | 2050 | D-003124-08 | GCCAATAGCCACTCTAAC   | € |
| EPHB6 |           |          |      |             |                      |   |
| EPHB6 | NM_004445 | 4758291  | 2051 | D-003125-05 | GGAAGTCGATCCTGCTTAT  | € |
| EPHB6 | NM_004445 | 4758291  | 2051 | D-003125-06 | GGACCAAGGTGGACACAAT  | € |
| EPHB6 | NM_004445 | 4758291  | 2051 | D-003125-07 | TGTGGGAAGTGATGAGTTA  | € |
| EPHB6 | NM_004445 | 4758291  | 2051 | D-003125-08 | CGGGAGACCTTCACCCTTT  | € |
| ERBB2 |           |          |      |             |                      |   |
| ERBB2 | NM_004448 | 4758297  | 2064 | D-003126-05 | GGACGAATTCTGCACAATG  | € |
| ERBB2 | NM_004448 | 4758297  | 2064 | D-003126-06 | GACGAATTCTGCACAATGG  | € |
| ERBB2 | NM_004448 | 4758297  | 2064 | D-003126-07 | CTACAACACAGACACGTTT  | € |
| ERBB2 | NM_004448 | 4758297  | 2064 | D-003126-08 | AGACGAAGCATACTGTGATG | € |
| ERBB3 |           |          |      |             |                      |   |
| ERBB3 | NM_001982 | 4503596  | 2065 | D-003127-05 | AAGAGGATGTCAACGGTTA  | € |
| ERBB3 | NM_001982 | 4503596  | 2065 | D-003127-06 | GAAGACTGCCAGACATTGA  | € |
| ERBB3 | NM_001982 | 4503596  | 2065 | D-003127-07 | GACAAACACTGGTGCTGAT  | € |
| ERBB3 | NM_001982 | 4503596  | 2065 | D-003127-08 | GCAGTGGATTGAGAAGTG   | € |
| ERBB4 |           |          |      |             |                      |   |
| ERBB4 | NM_005235 | 4885214  | 2066 | D-003128-05 | GAGGAAAGATGCCAATTAA  | € |
| ERBB4 | NM_005235 | 4885214  | 2066 | D-003128-06 | GCAGGAAACATCTATATTA  | € |
| ERBB4 | NM_005235 | 4885214  | 2066 | D-003128-07 | GATCACAACTGCTGCTTAA  | € |
| ERBB4 | NM_005235 | 4885214  | 2066 | D-003128-08 | CCTCAAAGATACTAGTTA   | € |
| FER   |           |          |      |             |                      |   |
| FER   | NM_005246 | 4885230  | 2241 | D-003129-05 | GGAGTGACCTGAAGAATT   | € |
| FER   | NM_005246 | 4885230  | 2241 | D-003129-06 | TAAAGCAGATTCCCATTAA  | € |
| FER   | NM_005246 | 4885230  | 2241 | D-003129-07 | GGAAAGTACTGTCCCAAATG | € |
| FER   | NM_005246 | 4885230  | 2241 | D-003129-08 | GAACAAACGGCTGCTAAAGA | € |
| FES   |           |          |      |             |                      |   |
| FES   | NM_002005 | 13376997 | 2242 | D-003130-05 | CGAGGATCCTGAAGCAGTA  | € |
| FES   | NM_002005 | 13376997 | 2242 | D-003130-06 | AGGAATACCTGGAGATTAG  | € |
| FES   | NM_002005 | 13376997 | 2242 | D-003130-07 | CAACAGGAGCTCCGGAATG  | € |
| FES   | NM_002005 | 13376997 | 2242 | D-003130-08 | GGTGTGGGTGAGCAGATT   | € |
| FGFR1 |           |          |      |             |                      |   |
| FGFR1 | NM_000604 | 13186232 | 2260 | D-003131-05 | TAAGAAATGTCTCCTTTGA  | € |
| FGFR1 | NM_000604 | 13186232 | 2260 | D-003131-06 | GAAGACTGCTGGAGTTAAT  | € |

|       |           |          |      |             |                      |   |
|-------|-----------|----------|------|-------------|----------------------|---|
| FGFR1 | NM_000604 | 13186232 | 2260 | D-003131-07 | GATGGTCCCTTGTATGTCA  | 7 |
| FGFR1 | NM_000604 | 13186232 | 2260 | D-003131-08 | CTTAAGAAATGTCTCCTT   | 7 |
| FGFR2 |           |          |      |             |                      |   |
| FGFR2 | NM_000141 | 13186239 | 2263 | D-003132-05 | CCAAATCTCTCAACCAGAA  | 7 |
| FGFR2 | NM_000141 | 13186239 | 2263 | D-003132-06 | GAACAGTATTACACCTAGTT | 7 |
| FGFR2 | NM_000141 | 13186239 | 2263 | D-003132-07 | GGCCAACACTGTCAAGTT   | 7 |
| FGFR2 | NM_000141 | 13186239 | 2263 | D-003132-08 | GTGAAGATGTTGAAAGATG  | 7 |
| FGFR3 |           |          |      |             |                      |   |
| FGFR3 | NM_000142 | 13112046 | 2261 | D-003133-05 | TGTCGGACCTGGTGTCTGA  | 7 |
| FGFR3 | NM_000142 | 13112046 | 2261 | D-003133-06 | GCATCAAGCTCGGGCATCA  | 7 |
| FGFR3 | NM_000142 | 13112046 | 2261 | D-003133-07 | GGACGGCACACCCTACGTT  | 7 |
| FGFR3 | NM_000142 | 13112046 | 2261 | D-003133-08 | TGCACAACCTCGACTACTA  | 7 |
| FGFR4 |           |          |      |             |                      |   |
| FGFR4 | NM_002011 | 13112051 | 2264 | D-003134-05 | GCACTGGAGTCTCGTGATG  | 7 |
| FGFR4 | NM_002011 | 13112051 | 2264 | D-003134-06 | CATAGGGACCTCTCGAATA  | 7 |
| FGFR4 | NM_002011 | 13112051 | 2264 | D-003134-07 | ATACGGACATCATCCTGTA  | 7 |
| FGFR4 | NM_002011 | 13112051 | 2264 | D-003134-08 | ATAGGGACCTCTCGAATAG  | 7 |
| FGR   |           |          |      |             |                      |   |
| FGR   | NM_005248 | 4885234  | 2268 | D-003135-05 | GCGATCATGTGAAGCATT   | 7 |
| FGR   | NM_005248 | 4885234  | 2268 | D-003135-06 | TCACTGAGCTCATCACCAA  | 7 |
| FGR   | NM_005248 | 4885234  | 2268 | D-003135-07 | GAAGAGTGGTACTTGGAA   | 7 |
| FGR   | NM_005248 | 4885234  | 2268 | D-003135-08 | CCCAGAAGCTGCCCTTT    | 7 |
| FLT1  |           |          |      |             |                      |   |
| FLT1  | NM_002019 | 4503748  | 2321 | D-003136-05 | GAGCAAACGTGACTTATT   | 7 |
| FLT1  | NM_002019 | 4503748  | 2321 | D-003136-06 | CCAAATGGGTTTCATGTTA  | 7 |
| FLT1  | NM_002019 | 4503748  | 2321 | D-003136-07 | CAACAAGGATGCAGCACTA  | 7 |
| FLT1  | NM_002019 | 4503748  | 2321 | D-003136-08 | GGACGTAACTGAAGAGGAT  | 7 |
| FLT3  |           |          |      |             |                      |   |
| FLT3  | NM_004119 | 4758395  | 2322 | D-003137-05 | GAAGGCATCTACACCATT   | 7 |
| FLT3  | NM_004119 | 4758395  | 2322 | D-003137-06 | GAAGGAGTCTGGAATAGAA  | 7 |
| FLT3  | NM_004119 | 4758395  | 2322 | D-003137-07 | GAATTAAGTCGTGTGTT    | 7 |
| FLT3  | NM_004119 | 4758395  | 2322 | D-003137-08 | GGAATTCACTCACTCTGA   | 7 |
| FLT4  |           |          |      |             |                      |   |
| FLT4  | NM_002020 | 4503752  | 2324 | D-003138-05 | GCAAGAACGTGCATCTGTT  | 7 |
| FLT4  | NM_002020 | 4503752  | 2324 | D-003138-06 | GCGAATACCTGTCCTACGA  | 7 |
| FLT4  | NM_002020 | 4503752  | 2324 | D-003138-07 | GAAGACATTGAGGAATT    | 7 |
| FLT4  | NM_002020 | 4503752  | 2324 | D-003138-08 | GAGCAGCCATTCACTCAACA | 7 |
| FRK   |           |          |      |             |                      |   |
| FRK   | NM_002031 | 4503786  | 2444 | D-003139-05 | GAAACAGACTCTTCATATT  | 7 |
| FRK   | NM_002031 | 4503786  | 2444 | D-003139-06 | GAACAATACCACTCCAGTA  | 7 |
| FRK   | NM_002031 | 4503786  | 2444 | D-003139-07 | CAAGACCGGTTCTTCTA    | 7 |
| FRK   | NM_002031 | 4503786  | 2444 | D-003139-08 | GCAAGAATATCTCCAAAAT  | 7 |
| FYN   |           |          |      |             |                      |   |
| FYN   | NM_002037 | 23510344 | 2534 | D-003140-05 | GGAATGGACTCATATGCAA  | 7 |
| FYN   | NM_002037 | 23510344 | 2534 | D-003140-06 | GCAGAAGAGTGGTACTTTG  | 7 |
| FYN   | NM_002037 | 23510344 | 2534 | D-003140-07 | CAAAGGAAGTTACTGGAT   | 7 |
| FYN   | NM_002037 | 23510344 | 2534 | D-003140-08 | GAAGAGTGGTACTTGGAA   | 7 |
| HCK   |           |          |      |             |                      |   |
| HCK   | NM_002110 | 4504356  | 3055 | D-003141-05 | GAGATAACCGTGAACATT   | 7 |
| HCK   | NM_002110 | 4504356  | 3055 | D-003141-06 | GCAGGGAGATACCGTAAA   | 7 |
| HCK   | NM_002110 | 4504356  | 3055 | D-003141-07 | CATCGTGGTTGCCCTGTAT  | 7 |
| HCK   | NM_002110 | 4504356  | 3055 | D-003141-08 | TGTGTAAGATTGCTGACTT  | 7 |
| ITK   |           |          |      |             |                      |   |
| ITK   | NM_005546 | 21614549 | 3702 | D-003144-05 | CAAATAATCTGGAAACCTA  | 7 |
| ITK   | NM_005546 | 21614549 | 3702 | D-003144-06 | GAAGAAACGAGGAATAATA  | 7 |
| ITK   | NM_005546 | 21614549 | 3702 | D-003144-07 | GAAACTCTCTCATCCAAA   | 7 |

|          |           |          |       |             |                     |   |
|----------|-----------|----------|-------|-------------|---------------------|---|
| ITK      | NM_005546 | 21614549 | 3702  | D-003144-08 | GGAATGGGCATGAAGGATA | 7 |
| JAK1     |           |          |       |             |                     |   |
| JAK1     | NM_002227 | 4504802  | 3716  | D-003145-05 | CCACATAGCTGATCTGAAA | 7 |
| JAK1     | NM_002227 | 4504802  | 3716  | D-003145-06 | TGAAATCACTCACATTGTA | 7 |
| JAK1     | NM_002227 | 4504802  | 3716  | D-003145-07 | TAAGGAACCTCTATCATGA | 7 |
| JAK1     | NM_002227 | 4504802  | 3716  | D-003145-08 | GCAGGTGGCTGTTAAATCT | 7 |
| JAK2     |           |          |       |             |                     |   |
| JAK2     | NM_004972 | 13325062 | 3717  | D-003146-05 | GCAAATAGATCCAGTTCTT | 7 |
| JAK2     | NM_004972 | 13325062 | 3717  | D-003146-06 | GAGCAAAGATCCAAGACTA | 7 |
| JAK2     | NM_004972 | 13325062 | 3717  | D-003146-07 | GCCAGAAACTGAAACTTA  | 7 |
| JAK2     | NM_004972 | 13325062 | 3717  | D-003146-08 | GTACAGATTCGCAGATT   | 7 |
| JAK3     |           |          |       |             |                     |   |
| JAK3     | NM_000215 | 4557680  | 3718  | D-003147-05 | GCGCCTATCTTCTCCTT   | 7 |
| JAK3     | NM_000215 | 4557680  | 3718  | D-003147-06 | CCAGAAATCGTAGACATTA | 7 |
| JAK3     | NM_000215 | 4557680  | 3718  | D-003147-07 | CCTCATCTCTCAGACTAT  | 7 |
| JAK3     | NM_000215 | 4557680  | 3718  | D-003147-08 | TGTACGAGCTCTCACCTA  | 7 |
| KDR      |           |          |       |             |                     |   |
| KDR      | NM_002253 | 11321596 | 3791  | D-003148-05 | GGAAATCTCTGCAAGCTA  | 7 |
| KDR      | NM_002253 | 11321596 | 3791  | D-003148-06 | GATTACAGATCTCCATTAA | 7 |
| KDR      | NM_002253 | 11321596 | 3791  | D-003148-07 | GCAGACAGATCTACGTTG  | 7 |
| KDR      | NM_002253 | 11321596 | 3791  | D-003148-08 | GCGATGGCCTCTCTGTAA  | 7 |
| KIAA1079 |           |          |       |             |                     |   |
| KIAA1079 | NM_014916 | 7662475  | 22853 | D-003149-05 | GAAATTCTCTCAACTGATG | 7 |
| KIAA1079 | NM_014916 | 7662475  | 22853 | D-003149-06 | GCAGAGGTCTTCACACTT  | 7 |
| KIAA1079 | NM_014916 | 7662475  | 22853 | D-003149-07 | TAAATGATCTTCAGACAGA | 7 |
| KIAA1079 | NM_014916 | 7662475  | 22853 | D-003149-08 | GAGCAGCCCTACTCTGATA | 7 |
| KIT      |           |          |       |             |                     |   |
| KIT      | NM_000222 | 4557694  | 3815  | D-003150-05 | AAACACGGCTTAAGCAATT | 7 |
| KIT      | NM_000222 | 4557694  | 3815  | D-003150-06 | GAACAGAACCTTCACTGAT | 7 |
| KIT      | NM_000222 | 4557694  | 3815  | D-003150-07 | GGGAAGCCCTCATGTCTGA | 7 |
| KIT      | NM_000222 | 4557694  | 3815  | D-003150-08 | GCAATTCCATTATGTGTT  | 7 |
| LCK      |           |          |       |             |                     |   |
| LCK      | NM_005356 | 20428651 | 3932  | D-003151-05 | GAACTGCCATTATCCCATA | 7 |
| LCK      | NM_005356 | 20428651 | 3932  | D-003151-06 | GAGAGGTGGTAAACATTA  | 7 |
| LCK      | NM_005356 | 20428651 | 3932  | D-003151-07 | GGGCCAAGTTCCCATTA   | 7 |
| LCK      | NM_005356 | 20428651 | 3932  | D-003151-08 | GCACGCTGCTCATCCGAAA | 7 |
| LTK      |           |          |       |             |                     |   |
| LTK      | NM_002344 | 4505044  | 4058  | D-003152-05 | TGAATTCACTCCTGCCAAT | 7 |
| LTK      | NM_002344 | 4505044  | 4058  | D-003152-06 | GTGGCAACCTCAACACTGA | 7 |
| LTK      | NM_002344 | 4505044  | 4058  | D-003152-07 | GGAGCTAGCTGTGGATAAC | 7 |
| LTK      | NM_002344 | 4505044  | 4058  | D-003152-08 | GCAAGTTGCCATCAGAA   | 7 |
| LYN      |           |          |       |             |                     |   |
| LYN      | NM_002350 | 4505054  | 4067  | D-003153-05 | GCAGATGGCTGTGCAGAA  | 7 |
| LYN      | NM_002350 | 4505054  | 4067  | D-003153-06 | GGAGAAGGCTTGTATTAGT | 7 |
| LYN      | NM_002350 | 4505054  | 4067  | D-003153-07 | GATGAGCTATGACATTA   | 7 |
| LYN      | NM_002350 | 4505054  | 4067  | D-003153-08 | GGTGCTAAGTTCCCTATTA | 7 |
| MATK     |           |          |       |             |                     |   |
| MATK     | NM_002378 | 21450841 | 4145  | D-003154-05 | TGAAGAATATCAAGTGTGA | 7 |
| MATK     | NM_002378 | 21450841 | 4145  | D-003154-06 | CCGCTCAGCTCCTGCAGTT | 7 |
| MATK     | NM_002378 | 21450841 | 4145  | D-003154-07 | TACTGAACCTGCAGCATT  | 7 |
| MATK     | NM_002378 | 21450841 | 4145  | D-003154-08 | TGGGAGGTCTCTCATATG  | 8 |
| MERTK    |           |          |       |             |                     |   |
| MERTK    | NM_006343 | 5453737  | 10461 | D-003155-05 | GAACTTACCTTACATAGCT | 8 |
| MERTK    | NM_006343 | 5453737  | 10461 | D-003155-06 | GGACCTGCATACTTACTTA | 8 |
| MERTK    | NM_006343 | 5453737  | 10461 | D-003155-07 | TGACAGGAATCTTCTAATT | 8 |
| MERTK    | NM_006343 | 5453737  | 10461 | D-003155-08 | GGTAATGGCTCAGTCATGA | 8 |

|        |           |          |      |             |                      |   |
|--------|-----------|----------|------|-------------|----------------------|---|
| MET    |           |          |      |             |                      |   |
| MET    | NM_000245 | 4557746  | 4233 | D-003156-05 | GAAAGAACCTCTAACATT   | 8 |
| MET    | NM_000245 | 4557746  | 4233 | D-003156-06 | GGACAAGGCTGACCATATG  | 8 |
| MET    | NM_000245 | 4557746  | 4233 | D-003156-07 | CCAATGACCTGCTGAAATT  | 8 |
| MET    | NM_000245 | 4557746  | 4233 | D-003156-08 | GAGCATAACATTAAACAAA  | 8 |
| MST1R  |           |          |      |             |                      |   |
| MST1R  | NM_002447 | 4505264  | 4486 | D-003157-05 | GGATGGAGCTGCTGGCTTT  | 8 |
| MST1R  | NM_002447 | 4505264  | 4486 | D-003157-06 | CTGCAGACCTATAGATTAA  | 8 |
| MST1R  | NM_002447 | 4505264  | 4486 | D-003157-07 | GCACCTGTCTCACTCTTGA  | 8 |
| MST1R  | NM_002447 | 4505264  | 4486 | D-003157-08 | GAAAGAGTCCATCCAGCTA  | 8 |
| MUSK   |           |          |      |             |                      |   |
| MUSK   | NM_005592 | 5031926  | 4593 | D-003158-05 | GAAGAAGCCTCGGCAGATA  | 8 |
| MUSK   | NM_005592 | 5031926  | 4593 | D-003158-06 | GTAATAATCTCCATCATGT  | 8 |
| MUSK   | NM_005592 | 5031926  | 4593 | D-003158-07 | GGAATGAACTGAAAGTAGT  | 8 |
| MUSK   | NM_005592 | 5031926  | 4593 | D-003158-08 | GAGATTCCTGGACTAGAA   | 8 |
| NTRK1  |           |          |      |             |                      |   |
| NTRK1  | NM_002529 | 4585711  | 4914 | D-003159-05 | GGACAACCCCTTCGAGTTC  | 8 |
| NTRK1  | NM_002529 | 4585711  | 4914 | D-003159-06 | CCAGTGACCTAACAGGAA   | 8 |
| NTRK1  | NM_002529 | 4585711  | 4914 | D-003159-07 | CCACAATACCTCAGTGATG  | 8 |
| NTRK1  | NM_002529 | 4585711  | 4914 | D-003159-08 | GAAGAGTGGCTCCGTTTC   | 8 |
| NTRK2  |           |          |      |             |                      |   |
| NTRK2  | NM_006180 | 21361305 | 4915 | D-003160-05 | GAACAGAAGTAATGAAATC  | 8 |
| NTRK2  | NM_006180 | 21361305 | 4915 | D-003160-06 | GTAATGCTGTTCTGCTTA   | 8 |
| NTRK2  | NM_006180 | 21361305 | 4915 | D-003160-07 | GCAAGACACTCCAAGTTG   | 8 |
| NTRK2  | NM_006180 | 21361305 | 4915 | D-003160-08 | GAAAGTCTATCACATTATC  | 8 |
| NTRK3  |           |          |      |             |                      |   |
| NTRK3  | NM_002530 | 4505474  | 4916 | D-003161-05 | GAGCGAATCTGCTAGTGAA  | 8 |
| NTRK3  | NM_002530 | 4505474  | 4916 | D-003161-06 | GAAGTTCACTACAGAGAGT  | 8 |
| NTRK3  | NM_002530 | 4505474  | 4916 | D-003161-07 | GGTCGACGGTCCAAATTTG  | 8 |
| NTRK3  | NM_002530 | 4505474  | 4916 | D-003161-08 | GAATATCACTCCATACAC   | 8 |
| PDGFRA |           |          |      |             |                      |   |
| PDGFRA | NM_006206 | 15451787 | 5156 | D-003162-05 | GAAACTCCTGGACTATTT   | 8 |
| PDGFRA | NM_006206 | 15451787 | 5156 | D-003162-06 | GAGATTTGGTCAACTATTT  | 8 |
| PDGFRA | NM_006206 | 15451787 | 5156 | D-003162-07 | GCACGCCGCTCCTGATAT   | 8 |
| PDGFRA | NM_006206 | 15451787 | 5156 | D-003162-08 | CATCAGAGCTGGATCTAGA  | 8 |
| PDGFRB |           |          |      |             |                      |   |
| PDGFRB | NM_002609 | 15451788 | 5159 | D-003163-05 | GAAAGGAGACGTCAAATAT  | 8 |
| PDGFRB | NM_002609 | 15451788 | 5159 | D-003163-06 | GGAATGAGGTGGTCAACTT  | 8 |
| PDGFRB | NM_002609 | 15451788 | 5159 | D-003163-07 | CAACGAGTCTCCAGTGCTA  | 8 |
| PDGFRB | NM_002609 | 15451788 | 5159 | D-003163-08 | GAGAGGACCTGCCGAGCAA  | 8 |
| PTK2   |           |          |      |             |                      |   |
| PTK2   | NM_005607 | 27886592 | 5747 | D-003164-05 | GAAGTTGGGTTGTCTAGAA  | 8 |
| PTK2   | NM_005607 | 27886592 | 5747 | D-003164-06 | GAAGAACAAATGATGTAATC | 8 |
| PTK2   | NM_005607 | 27886592 | 5747 | D-003164-07 | GGAAATTGCTTGAAGTTG   | 8 |
| PTK2   | NM_005607 | 27886592 | 5747 | D-003164-08 | GGTTCAAGCTGGATTATTT  | 8 |
| PTK2B  |           |          |      |             |                      |   |
| PTK2B  | NM_004103 | 27886583 | 2185 | D-003165-05 | GAACATGGCTGACCTCATA  | 8 |
| PTK2B  | NM_004103 | 27886583 | 2185 | D-003165-06 | GGACCACGCTGCTCTATTT  | 8 |
| PTK2B  | NM_004103 | 27886583 | 2185 | D-003165-07 | GGACGAGGACTATTACAAA  | 8 |
| PTK2B  | NM_004103 | 27886583 | 2185 | D-003165-08 | TGGCAGAGCTCATCAACAA  | 8 |
| PTK6   |           |          |      |             |                      |   |
| PTK6   | NM_005975 | 27886594 | 5753 | D-003166-05 | GAGAAAGTCCTGCCCGTTT  | 8 |
| PTK6   | NM_005975 | 27886594 | 5753 | D-003166-06 | TGAAGAAGCTGCGGGACAA  | 8 |
| PTK6   | NM_005975 | 27886594 | 5753 | D-003166-07 | CCGCGACTCTGATGAGAAA  | 8 |
| PTK6   | NM_005975 | 27886594 | 5753 | D-003166-08 | TGCCCGAGCTTGTGAACTA  | 8 |
| PTK7   |           |          |      |             |                      |   |

|       |           |          |       |             |                       |   |
|-------|-----------|----------|-------|-------------|-----------------------|---|
| PTK7  | NM_002821 | 27886610 | 5754  | D-003167-05 | GAGAGAAGCCCACTATTAA   | 8 |
| PTK7  | NM_002821 | 27886610 | 5754  | D-003167-06 | CGAGAGAAGCCCACTATTAA  | 8 |
| PTK7  | NM_002821 | 27886610 | 5754  | D-003167-07 | GGAGGGAGTTGGAGATGTT   | 8 |
| PTK7  | NM_002821 | 27886610 | 5754  | D-003167-08 | GAAGACATGCCGCTATTTG   | 8 |
| PTK9  |           |          |       |             |                       |   |
| PTK9  | NM_002822 | 4506274  | 5756  | D-003168-05 | GAAGAACTACGACAGATTAA  | 8 |
| PTK9  | NM_002822 | 4506274  | 5756  | D-003168-09 | GAAGGGAGACTATTTAGAGT  | 8 |
| PTK9  | NM_002822 | 4506274  | 5756  | D-003168-10 | GAGCGGATGCTGTATTCTA   | 8 |
| PTK9  | NM_002822 | 4506274  | 5756  | D-003168-11 | CTGCAGACTTCCTTTATGA   | 8 |
| PTK9L |           |          |       |             |                       |   |
| PTK9L | NM_007284 | 31543446 | 11344 | D-003169-05 | AGAGAGAGCTCCAGCAGAT   | 8 |
| PTK9L | NM_007284 | 31543446 | 11344 | D-003169-06 | TTAACGAGGTGAAGACAGA   | 8 |
| PTK9L | NM_007284 | 31543446 | 11344 | D-003169-07 | ACACAGAGCCCACGGATGT   | 8 |
| PTK9L | NM_007284 | 31543446 | 11344 | D-003169-08 | GCTGGGATCAGGACTATGA   | 8 |
| RET   |           |          |       |             |                       |   |
| RET   | NM_000323 | 21536316 | 5979  | D-003170-05 | GCAAAGACCTGGAGAAGAT   | 8 |
| RET   | NM_000323 | 21536316 | 5979  | D-003170-06 | GCACACGGCTGCATGAGAA   | 8 |
| RET   | NM_000323 | 21536316 | 5979  | D-003170-07 | GAACTGGCCTGGAGAGAGT   | 8 |
| RET   | NM_000323 | 21536316 | 5979  | D-003170-08 | TTAAATGGATGGCAATTGA   | 8 |
| ROR1  |           |          |       |             |                       |   |
| ROR1  | NM_005012 | 4826867  | 4919  | D-003171-05 | GCAAGCATTTACTAGGA     | 8 |
| ROR1  | NM_005012 | 4826867  | 4919  | D-003171-06 | GAGCAAGGCTAAAGAGCTA   | 8 |
| ROR1  | NM_005012 | 4826867  | 4919  | D-003171-07 | GAGAGCAACTCATGTAAA    | 8 |
| ROR1  | NM_005012 | 4826867  | 4919  | D-003171-08 | GAGAATGTCCGTGTCAAA    | 8 |
| ROR2  |           |          |       |             |                       |   |
| ROR2  | NM_004560 | 19743897 | 4920  | D-003172-05 | GGAACTCGCTGCTGCCTAT   | 8 |
| ROR2  | NM_004560 | 19743897 | 4920  | D-003172-06 | GCAGGTGCCCTCCTCAGATG  | 8 |
| ROR2  | NM_004560 | 19743897 | 4920  | D-003172-07 | GCAATGTGCTAGTGTACGA   | 8 |
| ROR2  | NM_004560 | 19743897 | 4920  | D-003172-08 | GAAGACAGAATATGGTTCA   | 8 |
| ROS1  |           |          |       |             |                       |   |
| ROS1  | NM_002944 | 19924164 | 6098  | D-003173-05 | GAGGAGACCTTCTTACTTA   | 8 |
| ROS1  | NM_002944 | 19924164 | 6098  | D-003173-06 | TTACAGAGGTTCAGGAGTA   | 8 |
| ROS1  | NM_002944 | 19924164 | 6098  | D-003173-07 | GAACAAACCTAACATGAA    | 8 |
| ROS1  | NM_002944 | 19924164 | 6098  | D-003173-08 | GAAAGAGCACTCAAATAA    | 8 |
| RYK   |           |          |       |             |                       |   |
| RYK   | NM_002958 | 11863158 | 6259  | D-003174-05 | GAAAGATGGTTACCGAATA   | 8 |
| RYK   | NM_002958 | 11863158 | 6259  | D-003174-06 | CAAAGTAGATTCTGAAGTT   | 8 |
| RYK   | NM_002958 | 11863158 | 6259  | D-003174-07 | TCACTACGCTCTATCCTTT   | 8 |
| RYK   | NM_002958 | 11863158 | 6259  | D-003174-08 | GGTGAAGGATATAGCAATA   | 8 |
| SRC   |           |          |       |             |                       |   |
| SRC   | NM_005417 | 21361210 | 6714  | D-003175-05 | GAGAACCTGGTGTGCAAAG   | 8 |
| SRC   | NM_005417 | 21361210 | 6714  | D-003175-09 | GAGAGAACCTGGTGTGCAA   | 8 |
| SRC   | NM_005417 | 21361210 | 6714  | D-003175-10 | GGAGTTTGCTGGACTTTCT   | 8 |
| SRC   | NM_005417 | 21361210 | 6714  | D-003175-11 | GAAAGTGAGACCACGAAAG   | 8 |
| SYK   |           |          |       |             |                       |   |
| SYK   | NM_003177 | 21361552 | 6850  | D-003176-05 | GGAATAATCTCAAGAATCA   | 8 |
| SYK   | NM_003177 | 21361552 | 6850  | D-003176-06 | GAACACTGGGCTCTGGTAATT | 8 |
| SYK   | NM_003177 | 21361552 | 6850  | D-003176-07 | GGAAGAACTGAGCAAATT    | 8 |
| SYK   | NM_003177 | 21361552 | 6850  | D-003176-08 | GAACAGACATGTCAAGGAT   | 8 |
| TEC   |           |          |       |             |                       |   |
| TEC   | NM_003215 | 4507428  | 7006  | D-003177-05 | GAAATTGTCTAGTAAGTGA   | 8 |
| TEC   | NM_003215 | 4507428  | 7006  | D-003177-06 | CACCTGAAGTGTAAATTAA   | 8 |
| TEC   | NM_003215 | 4507428  | 7006  | D-003177-07 | GTACAAAGTCGCAATCAA    | 8 |
| TEC   | NM_003215 | 4507428  | 7006  | D-003177-08 | TGGAGGAGATTCTTATTAA   | 8 |
| TEK   |           |          |       |             |                       |   |
| TEK   | NM_000459 | 4557868  | 7010  | D-003178-05 | GAAAGAAATATGCCTCCAAA  | 8 |

|       |           |          |      |             |                      |   |
|-------|-----------|----------|------|-------------|----------------------|---|
| TEK   | NM_000459 | 4557868  | 7010 | D-003178-06 | GGAATGACATCAAATTCA   | 8 |
| TEK   | NM_000459 | 4557868  | 7010 | D-003178-07 | TGAAGTACCTGATATTCTA  | 8 |
| TEK   | NM_000459 | 4557868  | 7010 | D-003178-08 | CGAAAGACCTACGTGAATA  | 8 |
| TIE   |           |          |      |             |                      |   |
| TIE   | NM_005424 | 4885630  | 7075 | D-003179-05 | GAGAGGAGGTTTATGTGAA  | 8 |
| TIE   | NM_005424 | 4885630  | 7075 | D-003179-06 | GGGACAGCCTCTACCCCTTA | 8 |
| TIE   | NM_005424 | 4885630  | 7075 | D-003179-07 | GAAGTTCTGTGCAAATTGG  | 8 |
| TIE   | NM_005424 | 4885630  | 7075 | D-003179-08 | CAACATGGCCTCAGAAGTG  | 9 |
| TNK1  |           |          |      |             |                      |   |
| TNK1  | NM_003985 | 4507610  | 8711 | D-003180-05 | GTTCTGGGCCTAAGTCTAA  | 9 |
| TNK1  | NM_003985 | 4507610  | 8711 | D-003180-06 | GAACCTGGGTCTACAAGATC | 9 |
| TNK1  | NM_003985 | 4507610  | 8711 | D-003180-07 | CGAGAGGTATCGGTATGA   | 9 |
| TNK1  | NM_003985 | 4507610  | 8711 | D-003180-08 | GGCGCATCCTGGAGCATT   | 9 |
| TXK   |           |          |      |             |                      |   |
| TXK   | NM_003328 | 4507742  | 7294 | D-003181-05 | GAACATCTATTGAGACAAG  | 9 |
| TXK   | NM_003328 | 4507742  | 7294 | D-003181-06 | TCAAGGCACCTTATGATT   | 9 |
| TXK   | NM_003328 | 4507742  | 7294 | D-003181-07 | GGAGAGGAATGGCTATATT  | 9 |
| TXK   | NM_003328 | 4507742  | 7294 | D-003181-08 | GGATATATGTGAAGGAATG  | 9 |
| TYK2  |           |          |      |             |                      |   |
| TYK2  | NM_003331 | 4507748  | 7297 | D-003182-05 | GAGGAGATCCACCACTTTA  | 9 |
| TYK2  | NM_003331 | 4507748  | 7297 | D-003182-06 | GCATCCACATTGCACATAA  | 9 |
| TYK2  | NM_003331 | 4507748  | 7297 | D-003182-07 | TCAAATACCTAGGCCACACT | 9 |
| TYK2  | NM_003331 | 4507748  | 7297 | D-003182-08 | CAATCTGCTGACGTCTG    | 9 |
| TYRO3 |           |          |      |             |                      |   |
| TYRO3 | NM_006293 | 27597077 | 7301 | D-003183-05 | GGTAGAAGGTGTGCCATT   | 9 |
| TYRO3 | NM_006293 | 27597077 | 7301 | D-003183-06 | ACGCTGAGATTACAACATA  | 9 |
| TYRO3 | NM_006293 | 27597077 | 7301 | D-003183-07 | GGATGGCTCCTTGTGAAA   | 9 |
| TYRO3 | NM_006293 | 27597077 | 7301 | D-003183-08 | GAGAGGAACTACGAAGATC  | 9 |
| YES1  |           |          |      |             |                      |   |
| YES1  | NM_005433 | 21071041 | 7525 | D-003184-05 | GAAGGACCTGATGAAAGA   | 9 |
| YES1  | NM_005433 | 21071041 | 7525 | D-003184-06 | TAAGAAGGGTGAAGATT    | 9 |
| YES1  | NM_005433 | 21071041 | 7525 | D-003184-07 | TCAAGAAGCTCAGATAATG  | 9 |
| YES1  | NM_005433 | 21071041 | 7525 | D-003184-08 | CAGAATCCCTCCATGAATT  | 9 |

Table VIII

5

| Gene Name | Acc#      | GI       | Locus Link | Duplex #    | Full Sequence        | SEQ. NO. |
|-----------|-----------|----------|------------|-------------|----------------------|----------|
| APC2      |           |          |            |             |                      |          |
| APC2      | NM_013366 | 7549800  | 29882      | D-003200-05 | GCAAGGACCTCTTCATCAA  | 921      |
| APC2      | NM_013366 | 7549800  | 29882      | D-003200-06 | GAGAAGAAGTCCACACTAT  | 922      |
| APC2      | NM_013366 | 7549800  | 29882      | D-003200-07 | GGAATGCCATCTCCCAATG  | 923      |
| APC2      | NM_013366 | 7549800  | 29882      | D-003200-09 | CAACACGTGTGACATCATC  | 924      |
| ATM       |           |          |            |             |                      |          |
| ATM       | NM_000051 | 20336202 | 472        | D-003201-05 | GCAAGCAGCTGAAACAAAT  | 925      |
| ATM       | NM_000051 | 20336202 | 472        | D-003201-06 | GAATGTTGCTTCTGAATT   | 926      |
| ATM       | NM_000051 | 20336202 | 472        | D-003201-07 | GACCTGAAGTCTTATTAA   | 927      |
| ATM       | NM_000051 | 20336202 | 472        | D-003201-08 | AGACAGAATTCCCAAATAA  | 928      |
| ATR       |           |          |            |             |                      |          |
| ATR       | NM_001184 | 20143978 | 545        | D-003202-05 | GAACAAACACTGCTGGTTG  | 929      |
| ATR       | NM_001184 | 20143978 | 545        | D-003202-06 | GAAGTCATCTGTTCAATT   | 930      |
| ATR       | NM_001184 | 20143978 | 545        | D-003202-07 | GAAATAAGGTAGACTCAAT  | 931      |
| ATR       | NM_001184 | 20143978 | 545        | D-003202-08 | CAACATAAAATCCAAGAAGA | 932      |
| BTAK      |           |          |            |             |                      |          |

|       |           |          |       |             |                       |     |
|-------|-----------|----------|-------|-------------|-----------------------|-----|
| BTAK  | NM_003600 | 3213196  | 6790  | D-003545-04 | CAAAGAACATCAGCTAGCAAA | 933 |
| BTAK  | NM_003600 | 3213196  | 6790  | D-003203-05 | GAAGAGAGTTATTCTATAGA  | 934 |
| BTAK  | NM_003600 | 3213196  | 6790  | D-003203-07 | CAAATGCCCTGTCTTACTG   | 935 |
| STK6  | NM_003600 | 3213196  | 6790  | D-003203-09 | TCTCGTGACTCAGCAAATT   | 936 |
| CCNA1 |           |          |       |             |                       |     |
| CCNA1 | NM_003914 | 16306528 | 8900  | D-003204-05 | GAACCTGGCTAACGTACGTA  | 937 |
| CCNA1 | NM_003914 | 16306528 | 8900  | D-003204-06 | GCAGATCCATTCTTGAAT    | 938 |
| CCNA1 | NM_003914 | 16306528 | 8900  | D-003204-07 | TCACAAGAACATCAGGTGTTA | 939 |
| CCNA1 | NM_003914 | 16306528 | 8900  | D-003204-08 | CATAAAGCGTACCTTGATA   | 940 |
| CCNA2 |           |          |       |             |                       |     |
| CCNA2 | NM_001237 | 16950653 | 890   | D-003205-05 | GCTGTGAACTACATTGATA   | 941 |
| CCNA2 | NM_001237 | 16950653 | 890   | D-003205-06 | GATGATACCTACACCAAGA   | 942 |
| CCNA2 | NM_001237 | 16950653 | 890   | D-003205-07 | GCTGTTAGCCTCAAAGTTT   | 943 |
| CCNA2 | NM_001237 | 16950653 | 890   | D-003205-08 | AAGCTGGCCTGAATCATT    | 944 |
| CCNB1 |           |          |       |             |                       |     |
| CCNB1 | NM_031966 | 14327895 | 891   | D-003206-05 | CAACATTACCTGTCAATATA  | 945 |
| CCNB1 | NM_031966 | 14327895 | 891   | D-003206-06 | CCAAATACCTGATGGAAC    | 946 |
| CCNB1 | NM_031966 | 14327895 | 891   | D-003206-07 | GAAATGTACCCCTCAGAAA   | 947 |
| CCNB1 | NM_031966 | 14327895 | 891   | D-003206-08 | GCACCTGGCTAACATGTA    | 948 |
| CCNB2 |           |          |       |             |                       |     |
| CCNB2 | NM_004701 | 10938017 | 9133  | D-003207-05 | CAACAAATGTCAACAAACA   | 949 |
| CCNB2 | NM_004701 | 10938017 | 9133  | D-003207-06 | GCAGCAAACCTCCTGAAGAT  | 950 |
| CCNB2 | NM_004701 | 10938017 | 9133  | D-003207-07 | CCAGTGATTTGGAGAATAT   | 951 |
| CCNB2 | NM_004701 | 10938017 | 9133  | D-003207-08 | GTGACTACGTTAAGGATAT   | 952 |
| CCNB3 |           |          |       |             |                       |     |
| CCNB3 | NM_033031 | 14719419 | 85417 | D-003208-05 | TGAACAAACTGCTGACTTT   | 953 |
| CCNB3 | NM_033031 | 14719419 | 85417 | D-003208-06 | GCTAGCTGCTGCCCTCCTTA  | 954 |
| CCNB3 | NM_033031 | 14719419 | 85417 | D-003208-07 | CAACTCACCTCGTGTGGAT   | 955 |
| CCNB3 | NM_033031 | 14719419 | 85417 | D-003208-08 | GTGGATCTCTACCTAACATGA | 956 |
| CCNC  |           |          |       |             |                       |     |
| CCNC  | NM_005190 | 7382485  | 892   | D-003209-05 | GCAGAGCTCCCACATTTG    | 957 |
| CCNC  | NM_005190 | 7382485  | 892   | D-003209-06 | GGAGTAGTTCAAATACAA    | 958 |
| CCNC  | NM_005190 | 7382485  | 892   | D-003209-07 | GACCTTGCTCCAGTATGT    | 959 |
| CCNC  | NM_005190 | 7382485  | 892   | D-003209-08 | GAGATTCTATGCCAGGTAT   | 960 |
| CCND1 |           |          |       |             |                       |     |
| CCND1 | NM_053056 | 16950654 | 595   | D-003210-05 | TGAACAAAGCTAACGTGGAA  | 961 |
| CCND1 | NM_053056 | 16950654 | 595   | D-003210-06 | CCAGAGTGATCAAGTGTGA   | 962 |
| CCND1 | NM_053056 | 16950654 | 595   | D-003210-07 | GTTCGTGGCCTCTAACATG   | 963 |
| CCND1 | NM_053056 | 16950654 | 595   | D-003210-08 | CCGAGAACGCTGTGCATCTA  | 964 |
| CCND2 |           |          |       |             |                       |     |
| CCND2 | NM_001759 | 16950656 | 894   | D-003211-06 | TGAATTACCTGGACCCTTT   | 965 |
| CCND2 | NM_001759 | 16950656 | 894   | D-003211-07 | CGGAGAACGCTGTGCATTAA  | 966 |
| CCND2 | NM_001759 | 16950656 | 894   | D-003211-08 | CTACAGACGTGCCGGATAT   | 967 |
| CCND2 | NM_001759 | 16950656 | 894   | D-003211-09 | CAACACAGACGTGGATTGT   | 968 |
| CCND3 |           |          |       |             |                       |     |
| CCND3 | NM_001760 | 16950657 | 896   | D-003212-05 | GGACCTGGCTGCTGTGATT   | 969 |
| CCND3 | NM_001760 | 16950657 | 896   | D-003212-06 | GATTATACCTTGCCTGT     | 970 |
| CCND3 | NM_001760 | 16950657 | 896   | D-003212-07 | GACCAGCACTCCTACAGAT   | 971 |
| CCND3 | NM_001760 | 16950657 | 896   | D-003212-08 | TGCGGAAGATGCTGGCTTA   | 972 |
| CCNE1 |           |          |       |             |                       |     |
| CCNE1 | NM_001238 | 17318558 | 898   | D-003213-05 | GTACTGAGCTGGGCAAATA   | 973 |
| CCNE1 | NM_001238 | 17318558 | 898   | D-003213-06 | GGAAATCTATCCTCCAAAG   | 974 |
| CCNE1 | NM_001238 | 17318558 | 898   | D-003213-07 | GGAGGTGTGTGAAGTCTAT   | 975 |
| CCNE1 | NM_001238 | 17318558 | 898   | D-003213-08 | CTAAATGACTTACATGAAG   | 976 |
| CCNE2 |           |          |       |             |                       |     |
| CCNE2 | NM_057749 | 17318564 | 9134  | D-003214-05 | GGATGGAACTCATTATATT   | 977 |

|        |           |          |       |             |                     |      |
|--------|-----------|----------|-------|-------------|---------------------|------|
| CCNE2  | NM_057749 | 17318564 | 9134  | D-003214-06 | GCAGATATGTCATGACAA  | 978  |
| CCNE2  | NM_057749 | 17318564 | 9134  | D-003214-07 | CATAATATCCAGACACATA | 979  |
| CCNE2  | NM_057749 | 17318564 | 9134  | D-003214-08 | TAAGAAAGCCTCAGGTTG  | 980  |
| CCNF   |           |          |       |             |                     |      |
| CCNF   | NM_001761 | 4502620  | 899   | D-003215-05 | TCACAAAGCATCCATATTG | 981  |
| CCNF   | NM_001761 | 4502620  | 899   | D-003215-06 | GAAGTCATGTTACAGTGT  | 982  |
| CCNF   | NM_001761 | 4502620  | 899   | D-003215-07 | TAGCCTACCTCTACAATGA | 983  |
| CCNF   | NM_001761 | 4502620  | 899   | D-003215-08 | GCACCCGGTTATCAGTAA  | 984  |
| CCNG1  |           |          |       |             |                     |      |
| CCNG1  | NM_004060 | 8670528  | 900   | D-003216-05 | GATAATGGCCTCAGAATGA | 985  |
| CCNG1  | NM_004060 | 8670528  | 900   | D-003216-06 | GCACGGCAATTGAAGCATA | 986  |
| CCNG1  | NM_004060 | 8670528  | 900   | D-003216-07 | GGAATAGAATGTCTTCAGA | 987  |
| CCNG1  | NM_004060 | 8670528  | 900   | D-003216-08 | TAACTCACCTTCCAACAAT | 988  |
| CCNG2  |           |          |       |             |                     |      |
| CCNG2  | NM_004354 | 4757935  | 901   | D-003217-05 | GGAGAGAGTTGGTTCTAA  | 989  |
| CCNG2  | NM_004354 | 4757935  | 901   | D-003217-06 | GGTGAACACCTAAACATTG | 990  |
| CCNG2  | NM_004354 | 4757935  | 901   | D-003217-07 | GAAATACTGAGCCTTGATA | 991  |
| CCNG2  | NM_004354 | 4757935  | 901   | D-003217-08 | TGCCAAAGTTGAAGATTAA | 992  |
| CCNH   |           |          |       |             |                     |      |
| CCNH   | NM_001239 | 17738313 | 902   | D-003218-05 | GCTGATGACTTTCTTAATA | 993  |
| CCNH   | NM_001239 | 17738313 | 902   | D-003218-06 | CAACTTAATTCCACCTTA  | 994  |
| CCNH   | NM_001239 | 17738313 | 902   | D-003218-07 | ATACACACCTTCCCAAATT | 995  |
| CCNH   | NM_001239 | 17738313 | 902   | D-003218-08 | GCTATGAAGATGATGATTA | 996  |
| CCNI   |           |          |       |             |                     |      |
| CCNI   | NM_006835 | 17738314 | 10983 | D-003219-05 | GCAAGCAGACCTCTACTAA | 997  |
| CCNI   | NM_006835 | 17738314 | 10983 | D-003219-07 | TGAGAGAATTCCAGTACTA | 998  |
| CCNI   | NM_006835 | 17738314 | 10983 | D-003219-08 | GGAATCAAACGGCTCTATA | 999  |
| CCNI   | NM_006835 | 17738314 | 10983 | D-003219-09 | GAATTGGGATCTTCACACA | 1000 |
| CCNT1  |           |          |       |             |                     |      |
| CCNT1  | NM_001240 | 17978465 | 904   | D-003220-05 | TATCAACACTGCTATAGTA | 1001 |
| CCNT1  | NM_001240 | 17978465 | 904   | D-003220-06 | GAACAAACGTCTGGTGAT  | 1002 |
| CCNT1  | NM_001240 | 17978465 | 904   | D-003220-07 | GCACAAGACTCACCCATCT | 1003 |
| CCNT1  | NM_001240 | 17978465 | 904   | D-003220-08 | GCACAGACTTCTACTTCA  | 1004 |
| CCNT2A |           |          |       |             |                     |      |
| CCNT2A | NM_001241 | 17978467 | 905   | D-003221-05 | GCACAGACATCCTATTCA  | 1005 |
| CCNT2A | NM_001241 | 17978467 | 905   | D-003221-06 | GCAGGGACCTTCTATATCA | 1006 |
| CCNT2A | NM_001241 | 17978467 | 905   | D-003221-07 | GAACAGCTATATTACAGA  | 1007 |
| CCNT2A | NM_001241 | 17978467 | 905   | D-003221-09 | TTATATAGCTGCCAGGTA  | 1008 |
| CCNT2B |           |          |       |             |                     |      |
| CCNT2B | NM_058241 | 17978468 | 905   | D-003222-05 | GCACAGACATCCTATTCA  | 1009 |
| CCNT2B | NM_058241 | 17978468 | 905   | D-003222-06 | GCAGGGACCTTCTATATCA | 1010 |
| CCNT2B | NM_058241 | 17978468 | 905   | D-003222-07 | GAACAGCTATATTACAGA  | 1011 |
| CCNT2B | NM_058241 | 17978468 | 905   | D-003222-08 | GGTGAATGTACCCAGTTA  | 1012 |
| CDC16  |           |          |       |             |                     |      |
| CDC16  | NM_003903 | 14110370 | 8881  | D-003223-05 | GTAGATGGCTTGCAAGAGA | 1013 |
| CDC16  | NM_003903 | 14110370 | 8881  | D-003223-06 | TAAAGTAGCTCACTCTCT  | 1014 |
| CDC16  | NM_003903 | 14110370 | 8881  | D-003223-07 | GCTACAAGCTTACTTCTGT | 1015 |
| CDC16  | NM_003903 | 14110370 | 8881  | D-003223-08 | TGGAAGAGCCCATCAATAA | 1016 |
| CDC2   |           |          |       |             |                     |      |
| CDC2   | NM_033379 | 27886643 | 983   | D-003552-01 | GTACAGATCTCCAGAAGTA | 1017 |
| CDC2   | NM_033379 | 27886643 | 983   | D-003552-02 | GATCAACTCTTCAGGATT  | 1018 |
| CDC2   | NM_033379 | 27886643 | 983   | D-003552-03 | GGTTATATCTCATCTTGA  | 1019 |
| CDC2   | NM_033379 | 27886643 | 983   | D-003552-04 | GAACCTCGTCATCCAAATA | 1020 |
| CDC20  |           |          |       |             |                     |      |
| CDC20  | NM_001255 | 4557436  | 991   | D-003225-05 | GGGAATATATATCCTCTGT | 1021 |
| CDC20  | NM_001255 | 4557436  | 991   | D-003225-06 | GAAACGGCTTCGAAATATG | 1022 |

|        |           |          |       |             |                      |      |
|--------|-----------|----------|-------|-------------|----------------------|------|
| CDC20  | NM_001255 | 4557436  | 991   | D-003225-07 | GAAGACCTGCCGTTACATT  | 1023 |
| CDC20  | NM_001255 | 4557436  | 991   | D-003225-08 | CACCAGTGATCGACACATT  | 1024 |
| CDC25A |           |          |       |             |                      |      |
| CDC25A | NM_001789 | 4502704  | 993   | D-003226-05 | GAAATTATGGCATCTGTTT  | 1025 |
| CDC25A | NM_001789 | 4502704  | 993   | D-003226-06 | TACAAGGAGTTCTTATGA   | 1026 |
| CDC25A | NM_001789 | 4502704  | 993   | D-003226-07 | CCACGAGGACTTAAAGAA   | 1027 |
| CDC25A | NM_001789 | 4502704  | 993   | D-003226-08 | TGGGAAACATCAGGATT    | 1028 |
| CDC25B |           |          |       |             |                      |      |
| CDC25B | NM_004358 | 11641416 | 994   | D-003227-05 | GCAGATAACCCCTATGAATA | 1029 |
| CDC25B | NM_004358 | 11641416 | 994   | D-003227-06 | CTAGGTCGCTTCTCTGA    | 1030 |
| CDC25B | NM_004358 | 11641416 | 994   | D-003227-07 | GAGAGCTGATTGGAGATTA  | 1031 |
| CDC25B | NM_004358 | 11641416 | 994   | D-003227-08 | AAAAGGACCTCGTCATGTA  | 1032 |
| CDC25C |           |          |       |             |                      |      |
| CDC25C | NM_001790 | 12408659 | 995   | D-003228-05 | GAGCAGAAGTGGCCTATAT  | 1033 |
| CDC25C | NM_001790 | 12408659 | 995   | D-003228-06 | CAGAAGAGATTCAGATGA   | 1034 |
| CDC25C | NM_001790 | 12408659 | 995   | D-003228-07 | CCAGGGAGCCTTAAACTTA  | 1035 |
| CDC25C | NM_001790 | 12408659 | 995   | D-003228-08 | GAAACTTGGTGGACAGTGA  | 1036 |
| CDC27  |           |          |       |             |                      |      |
| CDC27  | NM_001256 | 16554576 | 996   | D-003229-06 | CATGCAAGCTGAAAGAATA  | 1037 |
| CDC27  | NM_001256 | 16554576 | 996   | D-003229-07 | CAACACAAGTACCTAATCA  | 1038 |
| CDC27  | NM_001256 | 16554576 | 996   | D-003229-08 | GGAGATGGATCCTATTAC   | 1039 |
| CDC27  | NM_001256 | 16554576 | 996   | D-003229-09 | GAAAAGCCATGATGATATT  | 1040 |
| CDC34  |           |          |       |             |                      |      |
| CDC34  | NM_004359 | 16357476 | 997   | D-003230-05 | GCTCAGACCTCTTCTACGA  | 1041 |
| CDC34  | NM_004359 | 16357476 | 997   | D-003230-06 | GGACGAGGGCGATCTATAC  | 1042 |
| CDC34  | NM_004359 | 16357476 | 997   | D-003230-07 | GATCAGGAGTACACAGACA  | 1043 |
| CDC34  | NM_004359 | 16357476 | 997   | D-003230-08 | TGAACGAGCCAACACCTT   | 1044 |
| CDC37  |           |          |       |             |                      |      |
| CDC37  | NM_007065 | 16357478 | 11140 | D-003231-05 | GCGAGGAGACAGCCAATTA  | 1045 |
| CDC37  | NM_007065 | 16357478 | 11140 | D-003231-06 | CACAAGACCTCGTGGAAA   | 1046 |
| CDC37  | NM_007065 | 16357478 | 11140 | D-003231-07 | ACAATCGTCATGCAATT    | 1047 |
| CDC37  | NM_007065 | 16357478 | 11140 | D-003231-08 | GAGGAGAAATGTGCACTCA  | 1048 |
| CDC45L |           |          |       |             |                      |      |
| CDC45L | NM_003504 | 34335230 | 8318  | D-003232-05 | GCACACGGATCTCCCTTGA  | 1049 |
| CDC45L | NM_003504 | 34335230 | 8318  | D-003232-06 | GCAAACACCTGCTCAAGTC  | 1050 |
| CDC45L | NM_003504 | 34335230 | 8318  | D-003232-07 | TGAAGAGTCTGCAAATAAA  | 1051 |
| CDC45L | NM_003504 | 34335230 | 8318  | D-003232-08 | GGACGTGGATGCTCTGT    | 1052 |
| CDC6   |           |          |       |             |                      |      |
| CDC6   | NM_001254 | 16357469 | 990   | D-003233-05 | GAACACAGCTGCCAGAT    | 1053 |
| CDC6   | NM_001254 | 16357469 | 990   | D-003233-06 | GAGCAGAGATGTCCACTGA  | 1054 |
| CDC6   | NM_001254 | 16357469 | 990   | D-003233-07 | GGAAATATCTTAGCTACTG  | 1055 |
| CDC6   | NM_001254 | 16357469 | 990   | D-003233-08 | GGACGAAGATTGGTATTG   | 1056 |
| CDC7   |           |          |       |             |                      |      |
| CDC7   | NM_003503 | 11038647 | 8317  | D-003234-05 | GGAATGAGGTACCTGATGA  | 1057 |
| CDC7   | NM_003503 | 11038647 | 8317  | D-003234-06 | CAGGAAAGGTGTTACAAA   | 1058 |
| CDC7   | NM_003503 | 11038647 | 8317  | D-003234-07 | CTACACAAATGCACAAATT  | 1059 |
| CDC7   | NM_003503 | 11038647 | 8317  | D-003234-08 | GTACGGGAATATATGCTTA  | 1060 |
| CDK10  |           |          |       |             |                      |      |
| CDK10  | NM_003674 | 32528262 | 8558  | D-003235-05 | GAACGTGTTGGAACCA     | 1061 |
| CDK10  | NM_003674 | 32528262 | 8558  | D-003235-06 | GGAAGCAGCCCTACAACAA  | 1062 |
| CDK10  | NM_003674 | 32528262 | 8558  | D-003235-07 | GCACGCCAGTGAGAACAT   | 1063 |
| CDK10  | NM_003674 | 32528262 | 8558  | D-003235-08 | GGAAGCAGCCCTACAACAA  | 1064 |
| CDK2   |           |          |       |             |                      |      |
| CDK2   | NM_001798 | 16936527 | 1017  | D-003236-05 | GAGCTTAACCATCCTAATA  | 1065 |
| CDK2   | NM_001798 | 16936527 | 1017  | D-003236-06 | GAGCTTAACCATCCTAATA  | 1066 |
| CDK2   | NM_001798 | 16936527 | 1017  | D-003236-07 | GTACCGAGCTCCTGAAATC  | 1067 |

|        |           |          |      |             |                      |      |
|--------|-----------|----------|------|-------------|----------------------|------|
| CDK2   | NM_001798 | 16936527 | 1017 | D-003236-08 | GAGAGGTGGTGGCGCTTAA  | 1068 |
| CDK3   |           |          |      |             |                      |      |
| CDK3   | NM_001258 | 4557438  | 1018 | D-003237-05 | GAGCATTGGTTCATCTTT   | 1069 |
| CDK3   | NM_001258 | 4557438  | 1018 | D-003237-06 | GATCGGAGAGGGCACCTAT  | 1070 |
| CDK3   | NM_001258 | 4557438  | 1018 | D-003237-07 | GAAGCTCTATCTGGTGT    | 1071 |
| CDK3   | NM_001258 | 4557438  | 1018 | D-003237-08 | GCAGAGATGGTGAACCGAA  | 1072 |
| CDK4   |           |          |      |             |                      |      |
| CDK4   | NM_000075 | 456426   | 1019 | D-003238-05 | GCAGCACTCTTATCTACAT  | 1073 |
| CDK4   | NM_000075 | 456426   | 1019 | D-003238-06 | GGAGGAGGCCTCCCATCA   | 1074 |
| CDK4   | NM_000075 | 456426   | 1019 | D-003238-07 | TCGAAAGCCTCTCTTCTGT  | 1075 |
| CDK4   | NM_000075 | 456426   | 1019 | D-003238-08 | GTACCGAGCTCCCGAAGTT  | 1076 |
| CDK5   |           |          |      |             |                      |      |
| CDK5   | NM_004935 | 4826674  | 1020 | D-003239-05 | TGACCAAGCTGCCAGACTA  | 1077 |
| CDK5   | NM_004935 | 4826674  | 1020 | D-003239-06 | GAGCTGAAATTGGCTGATT  | 1078 |
| CDK5   | NM_004935 | 4826674  | 1020 | D-003239-07 | CAACATCCCTGGTGAACGT  | 1079 |
| CDK5   | NM_004935 | 4826674  | 1020 | D-003239-08 | GGATTCCCGTCCGCTGTTA  | 1080 |
| CDK6   |           |          |      |             |                      |      |
| CDK6   | NM_001259 | 16950658 | 1021 | D-003240-05 | GCAAAGACCTACTTCTGAA  | 1081 |
| CDK6   | NM_001259 | 16950658 | 1021 | D-003240-06 | GAAGAAGACTGGCCTAGAG  | 1082 |
| CDK6   | NM_001259 | 16950658 | 1021 | D-003240-07 | GGTCTGGACTTTCTTCATT  | 1083 |
| CDK6   | NM_001259 | 16950658 | 1021 | D-003240-08 | TAACAGATATCGATGAAC   | 1084 |
| CDK7   |           |          |      |             |                      |      |
| CDK7   | NM_001799 | 16950659 | 1022 | D-003241-05 | GGACATAGATCAGAAGCTA  | 1085 |
| CDK7   | NM_001799 | 16950659 | 1022 | D-003241-06 | CAATAGAGCTTATACACAT  | 1086 |
| CDK7   | NM_001799 | 16950659 | 1022 | D-003241-07 | CATACAAGGCTTATTCTTA  | 1087 |
| CDK7   | NM_001799 | 16950659 | 1022 | D-003241-08 | GGAGACGACTTACTAGATC  | 1088 |
| CDK8   |           |          |      |             |                      |      |
| CDK8   | NM_001260 | 4502744  | 1024 | D-003242-05 | CCACAGTACTCACATCAGA  | 1089 |
| CDK8   | NM_001260 | 4502744  | 1024 | D-003242-06 | GCAATAACCACACTAATGG  | 1090 |
| CDK8   | NM_001260 | 4502744  | 1024 | D-003242-07 | GAAGAAAGTGAGAGTTGTT  | 1091 |
| CDK8   | NM_001260 | 4502744  | 1024 | D-003242-08 | GAACATGACCTCTGGCATA  | 1092 |
| CDK9   |           |          |      |             |                      |      |
| CDK9   | NM_001261 | 17017983 | 1025 | D-003243-05 | GGCCAAACGTGGACAAC    | 1093 |
| CDK9   | NM_001261 | 17017983 | 1025 | D-003243-06 | TGACGTCCATGTTGAGTA   | 1094 |
| CDK9   | NM_001261 | 17017983 | 1025 | D-003243-07 | CCAACCAGACGGAGTTGA   | 1095 |
| CDK9   | NM_001261 | 17017983 | 1025 | D-003243-08 | GAAGGTGGCTCTGAAGAAG  | 1096 |
| CDKN1C |           |          |      |             |                      |      |
| CDKN1C | NM_000076 | 4557440  | 1028 | D-003244-05 | GACCAGAACCGCTGGGATT  | 1097 |
| CDKN1C | NM_000076 | 4557440  | 1028 | D-003244-06 | GGACCGAACGTGGACAGCGA | 1098 |
| CDKN1C | NM_000076 | 4557440  | 1028 | D-003244-08 | GCAAGAGATCAGCGCCTGA  | 1099 |
| CDKN1C | NM_000076 | 4557440  | 1028 | D-003244-09 | CCGCTGGGATTACGACTTC  | 1100 |
| CDKN2B |           |          |      |             |                      |      |
| CDKN2B | NM_004936 | 17981693 | 1030 | D-003245-05 | GCGAGGAGAACAGGGCAT   | 1101 |
| CDKN2B | NM_004936 | 17981693 | 1030 | D-003245-06 | CCAACGGAGTCACCGTT    | 1102 |
| CDKN2B | NM_004936 | 17981693 | 1030 | D-003245-07 | CGATCCAGGTATGATGAT   | 1103 |
| CDKN2B | NM_004936 | 17981693 | 1030 | D-003245-08 | CCTGGAAGCCGGCGCGGAT  | 1104 |
| CDKN2C |           |          |      |             |                      |      |
| CDKN2C | NM_001262 | 17981697 | 1031 | D-003246-05 | GGACACCGCCTGTGATTG   | 1105 |
| CDKN2C | NM_001262 | 17981697 | 1031 | D-003246-06 | GCCAGGAGACTGCTACTTA  | 1106 |
| CDKN2C | NM_001262 | 17981697 | 1031 | D-003246-07 | TGAAAGACCGAACTGGTT   | 1107 |
| CDKN2C | NM_001262 | 17981697 | 1031 | D-003246-08 | GAACCTGCCCTGCACTTG   | 1108 |
| CDKN2D |           |          |      |             |                      |      |
| CDKN2D | NM_001800 | 17981700 | 1032 | D-003247-05 | TGGCAGTTCAAGAGGGTCA  | 1109 |
| CDKN2D | NM_001800 | 17981700 | 1032 | D-003247-06 | CTCAGGACCTCGTGGACAT  | 1110 |
| CDKN2D | NM_001800 | 17981700 | 1032 | D-003247-07 | TGAAGGTCTAGTGGAGCA   | 1111 |
| CDKN2D | NM_001800 | 17981700 | 1032 | D-003247-08 | AGACGGCGCTGCAGGTCA   | 1112 |

|        |           |          |       |             |                       |  |      |
|--------|-----------|----------|-------|-------------|-----------------------|--|------|
| CDT1   |           |          |       |             |                       |  |      |
| CDT1   | NM_030928 | 19923847 | 81620 | D-003248-05 | CCAAGGAGGCACAGAAGCA   |  | 1113 |
| CDT1   | NM_030928 | 19923847 | 81620 | D-003248-06 | GCTTCAACGTGGATGAAGT   |  | 1114 |
| CDT1   | NM_030928 | 19923847 | 81620 | D-003248-07 | TCTCCGGGCCAGAAGATAA   |  | 1115 |
| CDT1   | NM_030928 | 19923847 | 81620 | D-003248-08 | GCGCAATGTTGGCCAGATC   |  | 1116 |
| CENPA  |           |          |       |             |                       |  |      |
| CENPA  | NM_001809 | 4585861  | 1058  | D-003249-05 | GCACACACCTCTTGATAAG   |  | 1117 |
| CENPA  | NM_001809 | 4585861  | 1058  | D-003249-06 | GCAAGAGAAATATGTGTTA   |  | 1118 |
| CENPA  | NM_001809 | 4585861  | 1058  | D-003249-07 | TTACATGCAGGCCGAGTTA   |  | 1119 |
| CENPA  | NM_001809 | 4585861  | 1058  | D-003249-08 | GAGACAAGGTTGGCTAAAG   |  | 1120 |
| CENPB  |           |          |       |             |                       |  |      |
| CENPB  | NM_001810 | 26105977 | 1059  | D-003250-05 | GGACATAGCCGCCTGCTTT   |  | 1121 |
| CENPB  | NM_001810 | 26105977 | 1059  | D-003250-06 | GCACGATCCTGAAGAACAA   |  | 1122 |
| CENPB  | NM_001810 | 26105977 | 1059  | D-003250-07 | GGAGGAGGGTGTATGTTGAT  |  | 1123 |
| CENPB  | NM_001810 | 26105977 | 1059  | D-003250-08 | CCGAATGGCTGCAGAGTCT   |  | 1124 |
| CENPC1 |           |          |       |             |                       |  |      |
| CENPC1 | NM_001812 | 4502778  | 1060  | D-003251-05 | GCGAATAGATTATCAAGGA   |  | 1125 |
| CENPC1 | NM_001812 | 4502778  | 1060  | D-003251-06 | GAACAGAACATCCATCACAAA |  | 1126 |
| CENPC1 | NM_001812 | 4502778  | 1060  | D-003251-07 | CCATAAACCTCACCCAGTA   |  | 1127 |
| CENPC1 | NM_001812 | 4502778  | 1060  | D-003251-08 | CAAGAGAACACGTTGAAA    |  | 1128 |
| CENPE  |           |          |       |             |                       |  |      |
| CENPE  | NM_001813 | 4502780  | 1062  | D-003252-05 | GAAGACAGCTCAAATAATA   |  | 1129 |
| CENPE  | NM_001813 | 4502780  | 1062  | D-003252-06 | CAACAAAGCTACTAAATCA   |  | 1130 |
| CENPE  | NM_001813 | 4502780  | 1062  | D-003252-07 | GGAAAGAAGTGCTACCATA   |  | 1131 |
| CENPE  | NM_001813 | 4502780  | 1062  | D-003252-08 | GGAAAGAAATGACACAGTT   |  | 1132 |
| CENPF  |           |          |       |             |                       |  |      |
| CENPF  | NM_016343 | 14670380 | 1063  | D-003253-05 | GCGAATATCTGAATTAGAA   |  | 1133 |
| CENPF  | NM_016343 | 14670380 | 1063  | D-003253-06 | GGAAATTAAATGCATCCTTA  |  | 1134 |
| CENPF  | NM_016343 | 14670380 | 1063  | D-003253-07 | GAGCGAGGCTGGTGGTTA    |  | 1135 |
| CENPF  | NM_016343 | 14670380 | 1063  | D-003253-08 | CAAGTCATCTTCATCTAA    |  | 1136 |
| CENPH  |           |          |       |             |                       |  |      |
| CENPH  | NM_022909 | 21264590 | 64946 | D-003254-05 | GAAAGAAGAGATTGCAATT   |  | 1137 |
| CENPH  | NM_022909 | 21264590 | 64946 | D-003254-06 | CAGAACAAATTATGCAAGA   |  | 1138 |
| CENPH  | NM_022909 | 21264590 | 64946 | D-003254-07 | CTAGTGTGCTCATGGATAA   |  | 1139 |
| CENPH  | NM_022909 | 21264590 | 64946 | D-003254-08 | GAAACACCTATTAGAGCTA   |  | 1140 |
| CHEK1  |           |          |       |             |                       |  |      |
| CHEK1  | NM_001274 | 20127419 | 1111  | D-003255-05 | CAAATTGGATGCAGACAAA   |  | 1141 |
| CHEK1  | NM_001274 | 20127419 | 1111  | D-003255-06 | GCAACAGTATTCGGTATA    |  | 1142 |
| CHEK1  | NM_001274 | 20127419 | 1111  | D-003255-07 | GGACTTCTCTCCAGTAAAC   |  | 1143 |
| CHEK1  | NM_001274 | 20127419 | 1111  | D-003255-08 | AAAGATAGATGGTACAACA   |  | 1144 |
| CHEK2  |           |          |       |             |                       |  |      |
| CHEK2  | NM_007194 | 22209010 | 11200 | D-003256-02 | CTCTTACATTGCATACATA   |  | 1145 |
| CHEK2  | NM_007194 | 22209010 | 11200 | D-003256-03 | TAAACGCCTGAAAAGAAGCT  |  | 1146 |
| CHEK2  | NM_007194 | 22209010 | 11200 | D-003256-04 | GCATAGGACTCAAGTGTCA   |  | 1147 |
| CHEK2  | NM_007194 | 22209010 | 11200 | D-003256-05 | GAAATTGCACTGTCACTAA   |  | 1148 |
| CNK    |           |          |       |             |                       |  |      |
| CNK    | NM_004073 | 4758015  | 1263  | D-003257-05 | GCGAGAACATCCTAAATGA   |  | 1149 |
| CNK    | NM_004073 | 4758015  | 1263  | D-003257-07 | GCAAGTGGGTTGACTACTC   |  | 1150 |
| CNK    | NM_004073 | 4758015  | 1263  | D-003257-08 | GCACATCCGTTGCCATCA    |  | 1151 |
| CNK    | NM_004073 | 4758015  | 1263  | D-003257-09 | GACCTCAAGTTGGGAAATT   |  | 1152 |
| CRI1   |           |          |       |             |                       |  |      |
| CRI1   | NM_014335 | 7656937  | 23741 | D-003258-05 | GTGATGAGATTATTGATAG   |  | 1153 |
| CRI1   | NM_014335 | 7656937  | 23741 | D-003258-06 | GGACGAGGGCGAGGAATT    |  | 1154 |
| CRI1   | NM_014335 | 7656937  | 23741 | D-003258-07 | GGAAACGGAGCCTTGCTAA   |  | 1155 |
| CRI1   | NM_014335 | 7656937  | 23741 | D-003258-08 | TCAATCGTCTGACCGAAGA   |  | 1156 |
| E2F1   |           |          |       |             |                       |  |      |

|              |           |          |       |             |                       |      |
|--------------|-----------|----------|-------|-------------|-----------------------|------|
| E2F1         | NM_005225 | 12669910 | 1869  | D-003259-05 | GAACAGGGCCACTGACTCT   | 1157 |
| E2F1         | NM_005225 | 12669910 | 1869  | D-003259-06 | TGGACCACCTGATGAATAT   | 1158 |
| E2F1         | NM_005225 | 12669910 | 1869  | D-003259-07 | CCCAGGAGGTCACTTCTGA   | 1159 |
| E2F1         | NM_005225 | 12669910 | 1869  | D-003259-08 | GGCTGGACCTGGAAACTGA   | 1160 |
| E2F2         |           |          |       |             |                       |      |
| E2F2         | NM_004091 | 34485718 | 1870  | D-003260-05 | GGGAGAAAGACTCGGTATGA  | 1161 |
| E2F2         | NM_004091 | 34485718 | 1870  | D-003260-06 | GAGGACAACCTGCAGATAT   | 1162 |
| E2F2         | NM_004091 | 34485718 | 1870  | D-003260-07 | TGAAGGAGCTGATGAACAC   | 1163 |
| E2F2         | NM_004091 | 34485718 | 1870  | D-003260-08 | CCAAGAAGTTCATTACCT    | 1164 |
| E2F3         |           |          |       |             |                       |      |
| E2F3         | NM_001949 | 12669913 | 1871  | D-003261-05 | GAAATTAGATGAAC TGATC  | 1165 |
| E2F3         | NM_001949 | 12669913 | 1871  | D-003261-06 | TGAAGTGCCTGACTCAATA   | 1166 |
| E2F3         | NM_001949 | 12669913 | 1871  | D-003261-07 | GAACAAGGCAGCAGAAGTG   | 1167 |
| E2F3         | NM_001949 | 12669913 | 1871  | D-003261-08 | GAAACACACAGTCCAATGA   | 1168 |
| E2F4         |           |          |       |             |                       |      |
| E2F4         | NM_001950 | 12669914 | 1874  | D-003262-05 | GGAGATTGCTGACAAACTG   | 1169 |
| E2F4         | NM_001950 | 12669914 | 1874  | D-003262-06 | GAAGGTATCGGGCTAATCG   | 1170 |
| E2F4         | NM_001950 | 12669914 | 1874  | D-003262-07 | GTGCAGAAGTCCAGGGAAT   | 1171 |
| E2F4         | NM_001950 | 12669914 | 1874  | D-003262-08 | GGACAGTGGTGAGCTCAGT   | 1172 |
| E2F5         |           |          |       |             |                       |      |
| E2F5         | NM_001951 | 12669916 | 1875  | D-003263-05 | GCAGATGACTACAAC TTTA  | 1173 |
| E2F5         | NM_001951 | 12669916 | 1875  | D-003263-06 | GACATCAGCTACAGATATA   | 1174 |
| E2F5         | NM_001951 | 12669916 | 1875  | D-003263-07 | CAACATGTCTCTGAAAGAA   | 1175 |
| E2F5         | NM_001951 | 12669916 | 1875  | D-003263-08 | GAAGACATCTGTAATTGCT   | 1176 |
| E2F6         |           |          |       |             |                       |      |
| E2F6         | NM_001952 | 12669917 | 1876  | D-003264-05 | TAAACAAAGGTTGCAACGAA  | 1177 |
| E2F6         | NM_001952 | 12669917 | 1876  | D-003264-06 | TAGCATATGTGACCTATCA   | 1178 |
| E2F6         | NM_001952 | 12669917 | 1876  | D-003264-07 | GAAACCAGATTGGATGTT    | 1179 |
| E2F6         | NM_001952 | 12669917 | 1876  | D-003264-09 | GGAAC TTTCTGACTTATCA  | 1180 |
| FOS          |           |          |       |             |                       |      |
| FOS          | NM_005252 | 6552332  | 2353  | D-003265-05 | GGGATAGCCTCTTACTA     | 1181 |
| FOS          | NM_005252 | 6552332  | 2353  | D-003265-06 | GAACAGTTATCTCCAGAAG   | 1182 |
| FOS          | NM_005252 | 6552332  | 2353  | D-003265-07 | GGAGACAGACCAACTAGAA   | 1183 |
| FOS          | NM_005252 | 6552332  | 2353  | D-003265-08 | AGACCGAGCCCTTGATGA    | 1184 |
| HIPK2        |           |          |       |             |                       |      |
| HIPK2        | NM_022740 | 13430859 | 28996 | D-003266-06 | GAGAAC TCACTCCAATCGAA | 1185 |
| HIPK2        | NM_022740 | 13430859 | 28996 | D-003266-07 | AGACAGGGATTAAGTCAAA   | 1186 |
| HIPK2        | NM_022740 | 13430859 | 28996 | D-003266-08 | GGACAAAGACAAC TAGGTT  | 1187 |
| HIPK2        | NM_022740 | 13430859 | 28996 | D-003266-09 | GCACACACGTCAAATCATG   | 1188 |
| HUS1         |           |          |       |             |                       |      |
| HUS1         | NM_004507 | 31077213 | 3364  | D-003267-05 | ACAAAGGCCTTATGCAATA   | 1189 |
| HUS1         | NM_004507 | 31077213 | 3364  | D-003267-06 | GAAGTGCACATAGATATTA   | 1190 |
| HUS1         | NM_004507 | 31077213 | 3364  | D-003267-07 | AAGCTTAAC TT CATCCTT  | 1191 |
| HUS1         | NM_004507 | 31077213 | 3364  | D-003267-08 | GAAC TCTTCAACGAATTT   | 1192 |
| JUN          |           |          |       |             |                       |      |
| JUN          | NM_002228 | 7710122  | 3725  | D-003268-05 | TGGAAACGACCTTCTATGA   | 1193 |
| JUN          | NM_002228 | 7710122  | 3725  | D-003268-06 | GAACTGCACAGCCAGAAC    | 1194 |
| JUN          | NM_002228 | 7710122  | 3725  | D-003268-07 | GAGCTGGAGCGCCTGATAA   | 1195 |
| JUN          | NM_002228 | 7710122  | 3725  | D-003268-08 | TAACGCAGCAGTTGCAAAC   | 1196 |
| JUNB         |           |          |       |             |                       |      |
| JUNB         | NM_002229 | 4504808  | 3726  | D-003269-05 | GCATCAAAGTGGAGCGCAA   | 1197 |
| JUNB         | NM_002229 | 4504808  | 3726  | D-003269-06 | TGGAAGACCAAGAGAGCGCAT | 1198 |
| JUNB         | NM_002229 | 4504808  | 3726  | D-003269-07 | CATACACAGCTACGGGATA   | 1199 |
| JUNB         | NM_002229 | 4504808  | 3726  | D-003269-08 | CCATCAACATGGAAGACCA   | 1200 |
| LOC510<br>53 |           |          |       |             |                       |      |

|        |           |          |       |             |                        |      |
|--------|-----------|----------|-------|-------------|------------------------|------|
| LOC510 | NM_015895 | 20127542 | 51053 | D-003270-05 | GGAGAAAGGCCTGTATGA     | 1201 |
| LOC510 | NM_015895 | 20127542 | 51053 | D-003270-06 | GAATAGTTCTGTCCCAAGA    | 1202 |
| LOC510 | NM_015895 | 20127542 | 51053 | D-003270-07 | GAACATGTACAGTATATGG    | 1203 |
| LOC510 | NM_015895 | 20127542 | 51053 | D-003270-08 | GCAGAAACAAGAAGAAATC    | 1204 |
| MAD2L1 |           |          |       |             |                        |      |
| MAD2L1 | NM_002358 | 6466452  | 4085  | D-003271-05 | GAAAGATGGCAGTTGATA     | 1205 |
| MAD2L1 | NM_002358 | 6466452  | 4085  | D-003271-06 | TAAATAATGTGGTGGAAACA   | 1206 |
| MAD2L1 | NM_002358 | 6466452  | 4085  | D-003271-07 | GAAATCCGTTCAGTGTATCA   | 1207 |
| MAD2L1 | NM_002358 | 6466452  | 4085  | D-003271-08 | TTACTCGAGTGCAGAAATA    | 1208 |
| MAD2L2 |           |          |       |             |                        |      |
| MAD2L2 | NM_006341 | 6006019  | 10459 | D-003272-05 | GGAAGAGCGCGCTCATAAA    | 1209 |
| MAD2L2 | NM_006341 | 6006019  | 10459 | D-003272-06 | TGGAAGAGCGCGCTCATAA    | 1210 |
| MAD2L2 | NM_006341 | 6006019  | 10459 | D-003272-07 | AGCCACTCCTGGAGAAGAA    | 1211 |
| MAD2L2 | NM_006341 | 6006019  | 10459 | D-003272-08 | TGGAGAAATTCTGTCTTTGA   | 1212 |
| MCM2   |           |          |       |             |                        |      |
| MCM2   | NM_004526 | 33356546 | 4171  | D-003273-05 | GAAGATCTTGCACAGCATT    | 1213 |
| MCM2   | NM_004526 | 33356546 | 4171  | D-003273-06 | GGATAAGGCTCGTCAGATC    | 1214 |
| MCM2   | NM_004526 | 33356546 | 4171  | D-003273-07 | CAGAGCAGGTGACATATCA    | 1215 |
| MCM2   | NM_004526 | 33356546 | 4171  | D-003273-08 | GCCGTGGCTCCTGTATGA     | 1216 |
| MCM3   |           |          |       |             |                        |      |
| MCM3   | NM_002388 | 33356548 | 4172  | D-003274-05 | GGACATCAATATTCTTCTA    | 1217 |
| MCM3   | NM_002388 | 33356548 | 4172  | D-003274-06 | GCCAGGACATCTCCAGTTA    | 1218 |
| MCM3   | NM_002388 | 33356548 | 4172  | D-003274-07 | GCAGGTATGACCAGTATAA    | 1219 |
| MCM3   | NM_002388 | 33356548 | 4172  | D-003274-08 | GGAAATGCCTCAAGTACAC    | 1220 |
| MCM4   |           |          |       |             |                        |      |
| MCM4   | XM_030274 | 22047061 | 4173  | D-003275-05 | GGACATATCTATTCTTACT    | 1221 |
| MCM4   | XM_030274 | 22047061 | 4173  | D-003275-06 | GATGTTAGTCACCACTGA     | 1222 |
| MCM4   | XM_030274 | 22047061 | 4173  | D-003275-07 | CCAGCTGCCTCATACTTTA    | 1223 |
| MCM4   | XM_030274 | 22047061 | 4173  | D-003275-08 | GAAAGTACAAGATCGGTAT    | 1224 |
| MCM5   |           |          |       |             |                        |      |
| MCM5   | NM_006739 | 23510447 | 4174  | D-003276-05 | GAAGATCCCTGGCATCATC    | 1225 |
| MCM5   | NM_006739 | 23510447 | 4174  | D-003276-06 | GAACAGGGTTACCATCATG    | 1226 |
| MCM5   | NM_006739 | 23510447 | 4174  | D-003276-07 | GGACAAACATTGACTTCATG   | 1227 |
| MCM5   | NM_006739 | 23510447 | 4174  | D-003276-08 | CCAAGGAGGTAGCTGATGA    | 1228 |
| MCM6   |           |          |       |             |                        |      |
| MCM6   | NM_005915 | 33469920 | 4175  | D-003277-05 | GGAAAGAGCTCAGAGATGA    | 1229 |
| MCM6   | NM_005915 | 33469920 | 4175  | D-003277-06 | GAGCAGCGATGGAGAAATT    | 1230 |
| MCM6   | NM_005915 | 33469920 | 4175  | D-003277-07 | GGAAACACCTGATGTCAAT    | 1231 |
| MCM6   | NM_005915 | 33469920 | 4175  | D-003277-08 | CAAACACATCTGCCGAAATC   | 1232 |
| MCM7   |           |          |       |             |                        |      |
| MCM7   | NM_005916 | 33469967 | 4176  | D-003278-05 | GGAAATATCCCTCGTAGTA    | 1233 |
| MCM7   | NM_005916 | 33469967 | 4176  | D-003278-06 | GGAAGAACGAGTCAGTCAAGTA | 1234 |
| MCM7   | NM_005916 | 33469967 | 4176  | D-003278-07 | CAACAAGCCTCGTGTGATC    | 1235 |
| MCM7   | NM_005916 | 33469967 | 4176  | D-003278-08 | GGAGAGAACACAAGGATTG    | 1236 |
| MDM2   |           |          |       |             |                        |      |
| MDM2   | NM_002392 | 4505136  | 4193  | D-003279-05 | GGAGATATGTTGTGAAAGA    | 1237 |
| MDM2   | NM_002392 | 4505136  | 4193  | D-003279-06 | CCACAAATCTGATAGTATT    | 1238 |
| MDM2   | NM_002392 | 4505136  | 4193  | D-003279-07 | GATGAGGTATATCAAGTTA    | 1239 |
| MDM2   | NM_002392 | 4505136  | 4193  | D-003279-08 | GGAAGAAACCCAAGACAAA    | 1240 |
| MKI67  |           |          |       |             |                        |      |
| MKI67  | NM_002417 | 19923216 | 4288  | D-003280-05 | GCACAAAGCTTGGTTATAA    | 1241 |
| MKI67  | NM_002417 | 19923216 | 4288  | D-003280-06 | CCTAAGACCTGAACATT      | 1242 |
| MKI67  | NM_002417 | 19923216 | 4288  | D-003280-07 | CAAAGAGGAACACAAATT     | 1243 |

|       |           |          |       |             |                       |      |
|-------|-----------|----------|-------|-------------|-----------------------|------|
| MKI67 | NM_002417 | 19923216 | 4288  | D-003280-08 | GTAAATGGGTCTGTTATTG   | 1244 |
| MNAT1 |           |          |       |             |                       |      |
| MNAT1 | NM_002431 | 4505224  | 4331  | D-003281-05 | GGAAGAACGTTAGAAGTG    | 1245 |
| MNAT1 | NM_002431 | 4505224  | 4331  | D-003281-06 | TAGATGAGCTGGAGAGTTC   | 1246 |
| MNAT1 | NM_002431 | 4505224  | 4331  | D-003281-07 | GGACCTGCTGGAGGCTAT    | 1247 |
| MNAT1 | NM_002431 | 4505224  | 4331  | D-003281-08 | GCAGATAGAGACATATGGA   | 1248 |
| MYC   |           |          |       |             |                       |      |
| MYC   | NM_002467 | 31543215 | 4609  | D-003282-05 | CAGAGAACGCTGGCCTCCTA  | 1249 |
| MYC   | NM_002467 | 31543215 | 4609  | D-003282-06 | GAAACGACGAGAACAGTTG   | 1250 |
| MYC   | NM_002467 | 31543215 | 4609  | D-003282-07 | CGACGAGACCTTCATCAAA   | 1251 |
| MYC   | NM_002467 | 31543215 | 4609  | D-003282-08 | CCACACATCAGCACAACTA   | 1252 |
| ORC1L |           |          |       |             |                       |      |
| ORC1L | NM_004153 | 31795543 | 4998  | D-003283-05 | GAACAGGAATTCCAAGACA   | 1253 |
| ORC1L | NM_004153 | 31795543 | 4998  | D-003283-06 | TAAGAACGTCGCTCGAGTA   | 1254 |
| ORC1L | NM_004153 | 31795543 | 4998  | D-003283-07 | GAGATCACCTCACCTTCTA   | 1255 |
| ORC1L | NM_004153 | 31795543 | 4998  | D-003283-08 | GCAGAGAGCCCTTCTTGGGA  | 1256 |
| ORC2L |           |          |       |             |                       |      |
| ORC2L | NM_006190 | 32454751 | 4999  | D-003284-05 | GAAGAACCTCCTATGAGA    | 1257 |
| ORC2L | NM_006190 | 32454751 | 4999  | D-003284-06 | GAAGGGAAC TGATGGAGTA  | 1258 |
| ORC2L | NM_006190 | 32454751 | 4999  | D-003284-07 | GAAGAATGATCCTGAGATT   | 1259 |
| ORC2L | NM_006190 | 32454751 | 4999  | D-003284-08 | GAAGAGATGTTCAAGAACATC | 1260 |
| ORC3L |           |          |       |             |                       |      |
| ORC3L | NM_012381 | 32483366 | 23595 | D-003285-05 | GGACTGCTGTAGATATA     | 1261 |
| ORC3L | NM_012381 | 32483366 | 23595 | D-003285-06 | GAAC TGATGACCATACTTG  | 1262 |
| ORC3L | NM_012381 | 32483366 | 23595 | D-003285-07 | AAAGATCTCTGCCAATA     | 1263 |
| ORC3L | NM_012381 | 32483366 | 23595 | D-003285-08 | CAGCACAGCTAACAGAGATA  | 1264 |
| ORC4L |           |          |       |             |                       |      |
| ORC4L | NM_002552 | 32454749 | 5000  | D-003286-06 | GAAAGCACATTCCGTTTAT   | 1265 |
| ORC4L | NM_002552 | 32454749 | 5000  | D-003286-07 | TGAAAGAAC TCGATGGAAAT | 1266 |
| ORC4L | NM_002552 | 32454749 | 5000  | D-003286-08 | GCTGAGAAC GTGGAATGAAA | 1267 |
| ORC4L | NM_002552 | 32454749 | 5000  | D-003286-09 | CCAGTGATCTTCAATTAG    | 1268 |
| ORC5L |           |          |       |             |                       |      |
| ORC5L | NM_002553 | 32454752 | 5001  | D-003287-05 | GAAATAACCTGTGAAACAT   | 1269 |
| ORC5L | NM_002553 | 32454752 | 5001  | D-003287-06 | CAGATTACCTCTCTAGTGA   | 1270 |
| ORC5L | NM_002553 | 32454752 | 5001  | D-003287-07 | GAAC TTCCATATTACTCTA  | 1271 |
| ORC5L | NM_002553 | 32454752 | 5001  | D-003287-08 | GTATTCAAGCTGATTCTAT   | 1272 |
| ORC6L |           |          |       |             |                       |      |
| ORC6L | NM_014321 | 32454755 | 23594 | D-003288-05 | GAACATGGCTTCAAAGATA   | 1273 |
| ORC6L | NM_014321 | 32454755 | 23594 | D-003288-06 | GGACAGGGCTTATTAATT    | 1274 |
| ORC6L | NM_014321 | 32454755 | 23594 | D-003288-07 | GAAAGAACATAGTGGTTGA   | 1275 |
| ORC6L | NM_014321 | 32454755 | 23594 | D-003288-08 | TATCAGAGCTGCTTAAAT    | 1276 |
| PCNA  |           |          |       |             |                       |      |
| PCNA  | NM_002592 | 33239449 | 5111  | D-003289-05 | GATCGAGGATGAAGAACAGGA | 1277 |
| PCNA  | NM_002592 | 33239449 | 5111  | D-003289-07 | GCCGAGATCTCAGCCATAT   | 1278 |
| PCNA  | NM_002592 | 33239449 | 5111  | D-003289-09 | GAGGCCTGCTGGGATATTA   | 1279 |
| PCNA  | NM_002592 | 33239449 | 5111  | D-003289-10 | GTGGAGAAC TTGGAAATGG  | 1280 |
| PLK   |           |          |       |             |                       |      |
| PLK   | NM_005030 | 21359872 | 5347  | D-003290-05 | CAACCAAAGTCGAATATGA   | 1281 |
| PLK   | NM_005030 | 21359872 | 5347  | D-003290-06 | CAAGAACATGAATACAGT    | 1282 |
| PLK   | NM_005030 | 21359872 | 5347  | D-003290-07 | GAAGATGTCCATGGAAATA   | 1283 |
| PLK   | NM_005030 | 21359872 | 5347  | D-003290-08 | CAACACGCCCTCATCCTCTA  | 1284 |
| PIN1  |           |          |       |             |                       |      |
| PIN1  | NM_006221 | 5453897  | 5300  | D-003291-05 | GGACCAAGGGAGGGAGGCCCT | 1285 |
| PIN1  | NM_006221 | 5453897  | 5300  | D-003291-06 | CGTCCTGGCGGGCAGGAGAA  | 1286 |
| PIN1  | NM_006221 | 5453897  | 5300  | D-003291-07 | CGGGAGAGGAGGACTTTGA   | 1287 |
| PIN1  | NM_006221 | 5453897  | 5300  | D-003291-08 | AGTCGGGAGAGGAGGACTT   | 1288 |

|       |           |          |      |             |                      |      |
|-------|-----------|----------|------|-------------|----------------------|------|
| PIN1L |           |          |      |             |                      |      |
| PIN1L | NM_006222 | 5453899  | 5301 | D-003292-06 | CGACCTGGCGGCAGGAAAT  | 1289 |
| PIN1L | NM_006222 | 5453899  | 5301 | D-003292-07 | AGGCAGGAGAGAAGGACTT  | 1290 |
| PIN1L | NM_006222 | 5453899  | 5301 | D-003292-08 | GCTACATCCAGAAAGATCAA | 1291 |
| PIN1L | NM_006222 | 5453899  | 5301 | D-003292-09 | GGACAGTGTTCACGGATTC  | 1292 |
| RAD1  |           |          |      |             |                      |      |
| RAD1  | NM_002853 | 19718797 | 5810 | D-003293-05 | GAAGATGGACAAATATGTT  | 1293 |
| RAD1  | NM_002853 | 19718797 | 5810 | D-003293-06 | GGAAGAGTCTGTTACTTTT  | 1294 |
| RAD1  | NM_002853 | 19718797 | 5810 | D-003293-07 | GATAACAGAGGCTTCCTTT  | 1295 |
| RAD1  | NM_002853 | 19718797 | 5810 | D-003293-08 | GCATTAGTCCTATCTTGTA  | 1296 |
| RAD17 |           |          |      |             |                      |      |
| RAD17 | NM_133338 | 19718783 | 5884 | D-003294-05 | GAATCAAGCTCCATATGT   | 1297 |
| RAD17 | NM_133338 | 19718783 | 5884 | D-003294-06 | CAACAAAGCCCGAGGATAT  | 1298 |
| RAD17 | NM_133338 | 19718783 | 5884 | D-003294-07 | ACACATGCCTGGAGACTTA  | 1299 |
| RAD17 | NM_133338 | 19718783 | 5884 | D-003294-08 | CTACATAGATTCTTCATG   | 1300 |
| RAD9A |           |          |      |             |                      |      |
| RAD9A | NM_004584 | 19924112 | 5883 | D-003295-05 | TCAGCAAACCTGAATCTTA  | 1301 |
| RAD9A | NM_004584 | 19924112 | 5883 | D-003295-06 | GACATTGACTCTTACATGA  | 1302 |
| RAD9A | NM_004584 | 19924112 | 5883 | D-003295-08 | GGAAACCAACTATAGGCAAT | 1303 |
| RAD9A | NM_004584 | 19924112 | 5883 | D-003295-09 | CGGACGACTTTGCCAATGA  | 1304 |
| RB1   |           |          |      |             |                      |      |
| RB1   | NM_000321 | 19924112 | 5925 | D-003296-05 | GAAAGGACATGTGAACCTTA | 1305 |
| RB1   | NM_000321 | 19924112 | 5925 | D-003296-06 | GAAGAAGTATGATGTATTG  | 1306 |
| RB1   | NM_000321 | 4506434  | 5925 | D-003296-07 | GAAATGACTTCTACTCGAA  | 1307 |
| RB1   | NM_000321 | 4506434  | 5925 | D-003296-08 | GGAGGGAACATCTATATT   | 1308 |
| RBBP2 |           |          |      |             |                      |      |
| RBBP2 | NM_005056 | 4826967  | 5927 | D-003297-05 | CAAAGAAAGCTGAATAAACT | 1309 |
| RBBP2 | NM_005056 | 4826967  | 5927 | D-003297-06 | CAACACATATGGCGGGATT  | 1310 |
| RBBP2 | NM_005056 | 4826967  | 5927 | D-003297-07 | GGACAAACCTAGAAAGAAG  | 1311 |
| RBBP2 | NM_005056 | 4826967  | 5927 | D-003297-08 | GAAAGGCACTCTCTGT     | 1312 |
| RBL1  |           |          |      |             |                      |      |
| RBL1  | NM_002895 | 34577078 | 5933 | D-003298-05 | CAAGAGAAGTTGTGGCATA  | 1313 |
| RBL1  | NM_002895 | 34577078 | 5933 | D-003298-06 | CAGCAGCACTCCATTATA   | 1314 |
| RBL1  | NM_002895 | 34577078 | 5933 | D-003298-07 | ACAGAAAGGTCTATCATTT  | 1315 |
| RBL1  | NM_002895 | 34577078 | 5933 | D-003298-08 | GGACATAAAGTTACAATT   | 1316 |
| RBL2  |           |          |      |             |                      |      |
| RBL2  | NM_005611 | 21361291 | 5934 | D-003299-05 | GAGCAGAGCTTAATCGAAT  | 1317 |
| RBL2  | NM_005611 | 21361291 | 5934 | D-003299-06 | GAGAATAGCCCTTGTGTGA  | 1318 |
| RBL2  | NM_005611 | 21361291 | 5934 | D-003299-07 | GGACTTAGTTATGGAAAT   | 1319 |
| RBL2  | NM_005611 | 21361291 | 5934 | D-003299-08 | GAATTAGATGAGCGGATA   | 1320 |
| RBP1  |           |          |      |             |                      |      |
| RBP1  | NM_002899 | 8400726  | 5947 | D-003300-05 | GAGACAAGCTCCAGTGTGT  | 1321 |
| RBP1  | NM_002899 | 8400726  | 5947 | D-003300-06 | GCAAGCAAGTATTCAAGAA  | 1322 |
| RBP1  | NM_002899 | 8400726  | 5947 | D-003300-07 | GCAGGACGGTGACCATATG  | 1323 |
| RBP1  | NM_002899 | 8400726  | 5947 | D-003300-08 | GCAAGTGCATGACAACAGT  | 1324 |
| RPA3  |           |          |      |             |                      |      |
| RPA3  | NM_002947 | 19923751 | 6119 | D-003322-05 | GGAAGTGGTGGAAAGAGTA  | 1325 |
| RPA3  | NM_002947 | 19923751 | 6119 | D-003322-06 | GAAGATAGCCATCCTTTG   | 1326 |
| RPA3  | NM_002947 | 19923751 | 6119 | D-003322-07 | CATGCTAGCTCAATTATC   | 1327 |
| RPA3  | NM_002947 | 19923751 | 6119 | D-003322-08 | GATCTGGACTTTACAATG   | 1328 |
| SKP1A |           |          |      |             |                      |      |
| SKP1A | NM_006930 | 25777710 | 6500 | D-003323-05 | GGAGAGATATTGAAGTTG   | 1329 |
| SKP1A | NM_006930 | 25777710 | 6500 | D-003323-06 | GGGAATGGATGATGAAGGA  | 1330 |
| SKP1A | NM_006930 | 25777710 | 6500 | D-003323-07 | CAAACAACTGTGACTATT   | 1331 |
| SKP1A | NM_006930 | 25777710 | 6500 | D-003323-08 | TCAATTAGTTGCAGAGTT   | 1332 |
| SKP2  |           |          |      |             |                      |      |

|       |           |          |       |             |                      |      |
|-------|-----------|----------|-------|-------------|----------------------|------|
| SKP2  | NM_005983 | 16306594 | 6502  | D-003324-05 | CATCTAGACTTAAGTGATA  | 1333 |
| SKP2  | NM_005983 | 16306594 | 6502  | D-003324-06 | GAAATCAGATCTCTCTACT  | 1334 |
| SKP2  | NM_005983 | 16306594 | 6502  | D-003324-07 | CTAAAGGTCTCTGGTGT    | 1335 |
| SKP2  | NM_005983 | 16306594 | 6502  | D-003324-08 | GATGGTACCCCTCAACTGT  | 1336 |
| SNK   |           |          |       |             |                      |      |
| SNK   | NM_006622 | 5730054  | 10769 | D-003325-05 | GAAGACATCTACAAGCTTA  | 1337 |
| SNK   | NM_006622 | 5730054  | 10769 | D-003325-06 | GAAATACCTTCATGAACAA  | 1338 |
| SNK   | NM_006622 | 5730054  | 10769 | D-003325-07 | GAAGGTCAATGGCTCATAT  | 1339 |
| SNK   | NM_006622 | 5730054  | 10769 | D-003325-08 | CCGGAGATCTGGGGATTA   | 1340 |
| STK12 |           |          |       |             |                      |      |
| STK12 | NM_004217 | 4759177  | 9212  | D-003326-07 | CAGAAGAGCTGCACATTG   | 1341 |
| STK12 | NM_004217 | 4759177  | 9212  | D-003326-08 | CCAAACTGCTCAGGCATAA  | 1342 |
| STK12 | NM_004217 | 4759177  | 9212  | D-003326-09 | ACGCGGCACTTCACAATTG  | 1343 |
| STK12 | NM_004217 | 4759177  | 9212  | D-003326-10 | TGGGACACCCGACATCTTA  | 1344 |
| TFDP1 |           |          |       |             |                      |      |
| TFDP1 | NM_007111 | 34147667 | 7027  | D-003327-05 | GGAAGCAGCTCTGCCAAA   | 1345 |
| TFDP1 | NM_007111 | 34147667 | 7027  | D-003327-06 | GAGGAGACTTGAAAGAATA  | 1346 |
| TFDP1 | NM_007111 | 34147667 | 7027  | D-003327-07 | GAACCTAGAGGTGGAAAGA  | 1347 |
| TFDP1 | NM_007111 | 34147667 | 7027  | D-003327-08 | GCGAGAAGGTGCAGAGGAA  | 1348 |
| TFDP2 |           |          |       |             |                      |      |
| TFDP2 | NM_006286 | 5454111  | 7029  | D-003328-05 | GAAAGTGTGTGAGAAAGTT  | 1349 |
| TFDP2 | NM_006286 | 5454111  | 7029  | D-003328-06 | CACAGGACCTTCTGGTTA   | 1350 |
| TFDP2 | NM_006286 | 5454111  | 7029  | D-003328-07 | CGAAATCCCTGGTGCCAAA  | 1351 |
| TFDP2 | NM_006286 | 5454111  | 7029  | D-003328-08 | TGAGATCCATGATGACATA  | 1352 |
| TP53  |           |          |       |             |                      |      |
| TP53  | NM_000546 | 8400737  | 7157  | D-003329-05 | GAGGTTGGCTCTGACTGTA  | 1353 |
| TP53  | NM_000546 | 8400737  | 7157  | D-003329-06 | CAGTCTACCTCCGCCATA   | 1354 |
| TP53  | NM_000546 | 8400737  | 7157  | D-003329-07 | GCACAGAGGAAGAGAAATCT | 1355 |
| TP53  | NM_000546 | 8400737  | 7157  | D-003329-08 | GAAGAAACCACTGGATGGA  | 1356 |
| TP63  |           |          |       |             |                      |      |
| TP63  | NM_003722 | 31543817 | 8626  | D-003330-05 | CATCATGTCTGGACTATT   | 1357 |
| TP63  | NM_003722 | 31543817 | 8626  | D-003330-06 | CAAACAAGATTGAGATTAG  | 1358 |
| TP63  | NM_003722 | 31543817 | 8626  | D-003330-07 | GCACACAGACAAATGAATT  | 1359 |
| TP63  | NM_003722 | 31543817 | 8626  | D-003330-08 | CGACAGTCTGTACAATT    | 1360 |
| TP73  |           |          |       |             |                      |      |
| TP73  | NM_005427 | 4885644  | 7161  | D-003331-05 | GCAAGCAGCCCATCAAGGA  | 1361 |
| TP73  | NM_005427 | 4885644  | 7161  | D-003331-06 | GAGACGAGGACACGTACTA  | 1362 |
| TP73  | NM_005427 | 4885644  | 7161  | D-003331-07 | CTGCAGAACCTGACCATTG  | 1363 |
| TP73  | NM_005427 | 4885644  | 7161  | D-003331-08 | GGCCATGCCCTGTTACAAG  | 1364 |
| YWHAZ |           |          |       |             |                      |      |
| YWHAZ | NM_003406 | 21735623 | 7534  | D-003332-05 | GCAAGGAGCTGAATTATCC  | 1365 |
| YWHAZ | NM_003406 | 21735623 | 7534  | D-003332-06 | TAAGAGATATCTGCAATGA  | 1366 |
| YWHAZ | NM_003406 | 21735623 | 7534  | D-003332-07 | GACGGAAGGTGCTGAGAAA  | 1367 |
| YWHAZ | NM_003406 | 21735623 | 7534  | D-003332-08 | AGAGCAAAGTCTTCTATT   | 1368 |

Table IX

| Gene Name | Accession # | GI#      | Duplex #    | Sequence            | SEQ. ID NO. |
|-----------|-------------|----------|-------------|---------------------|-------------|
| AR        | NM_000044   | 21322251 | D-003400-01 | GGAACTCGATCGTATCATT | 1369        |
| AR        | NM_000044   | 21322251 | D-003400-02 | CAAGGGAGGTACACCAA   | 1370        |
| AR        | NM_000044   | 21322251 | D-003400-03 | TCAAGGAACTCGATCGTAT | 1371        |
| AR        | NM_000044   | 21322251 | D-003400-04 | GAAATGATTGCACTATTGA | 1372        |
| ESR1      | NM_000125   | 4503602  | D-003401-01 | GAATGTGCCTGGCTAGAGA | 1373        |
| ESR1      | NM_000125   | 4503602  | D-003401-02 | CATGAGAGCTGCCAACCTT | 1374        |
| ESR1      | NM_000125   | 4503602  | D-003401-03 | AGAGAAAGATTGGCCAGTA | 1375        |

|          |           |          |             |                       |      |
|----------|-----------|----------|-------------|-----------------------|------|
| ESR1     | NM_000125 | 4503602  | D-003401-04 | CAAGGAGACTCGCTACTGT   | 1376 |
|          |           |          |             |                       |      |
| ESR2     | NM_001437 | 10835012 | D-003402-01 | GAACATCTGCTAACATGA    | 1377 |
| ESR2     | NM_001437 | 10835012 | D-003402-02 | GCACGGCTCCATATACATA   | 1378 |
| ESR2     | NM_001437 | 10835012 | D-003402-03 | CAAGAAGATTCCCGGCTT    | 1379 |
| ESR2     | NM_001437 | 10835012 | D-003402-04 | GGAAATGCGTAGAAGGAAT   | 1380 |
|          |           |          |             |                       |      |
| ESRR A   | NM_004451 | 18860919 | D-003403-01 | GGCCTTCGCTGAGGACTTA   | 1381 |
| ESRR A   | NM_004451 | 18860919 | D-003403-02 | TGAATGCACTGGTGTCTCA   | 1382 |
| ESRR A   | NM_004451 | 18860919 | D-003403-03 | GCATTGAGCCTCTACAT     | 1383 |
| ESRR A   | NM_004451 | 18860919 | D-003403-04 | CCAGACAGCGGGCAAAGTG   | 1384 |
|          |           |          |             |                       |      |
| ESRR B   | NM_004452 | 22035686 | D-003404-01 | TACCTGAGCTTACAAATT    | 1385 |
| ESRR B   | NM_004452 | 22035686 | D-003404-02 | GCACTTCTATAGCGTCAA    | 1386 |
| ESRR B   | NM_004452 | 22035686 | D-003404-03 | CAACTCCGATTCCATGTAC   | 1387 |
| ESRR B   | NM_004452 | 22035686 | D-003404-04 | GGACTCGCCACCCATGTT    | 1388 |
|          |           |          |             |                       |      |
| ESRR G   | NM_001438 | 4503604  | D-003405-01 | AAACAAAGATCGACACATT   | 1389 |
| ESRR G   | NM_001438 | 4503604  | D-003405-02 | TCAGGAAACTGTATGATGA   | 1390 |
| ESRR G   | NM_001438 | 4503604  | D-003405-03 | GAAGACCAGTCCAAATTAG   | 1391 |
| ESRR G   | NM_001438 | 4503604  | D-003405-04 | ATGAAGCGCTGCAGGATTA   | 1392 |
|          |           |          |             |                       |      |
| HNF4A    | NM_000457 | 21361184 | D-003406-01 | CGACATCACTGGAGCATAT   | 1393 |
| HNF4A    | NM_000457 | 21361184 | D-003406-02 | GAAGGAAGCCGTCCAGAAT   | 1394 |
| HNF4A    | NM_000457 | 21361184 | D-003406-03 | CCAAGTACATCCCAGCTT    | 1395 |
| HNF4A    | NM_000457 | 21361184 | D-003406-04 | GGACATGGCCGACTACAGT   | 1396 |
|          |           |          |             |                       |      |
| HNF4G    | NM_004133 | 6631087  | D-003407-01 | GCACTGACATAAACGTTAA   | 1397 |
| HNF4G    | NM_004133 | 6631087  | D-003407-02 | ACAAAGAGATCCATGATGT   | 1398 |
| HNF4G    | NM_004133 | 6631087  | D-003407-03 | AGAGATCCATGATGTATAA   | 1399 |
| HNF4G    | NM_004133 | 6631087  | D-003407-04 | AAATGAACGTGACAGAATA   | 1400 |
|          |           |          |             |                       |      |
| HSAJ2425 | NM_017532 | 8923776  | D-003408-01 | GAATGAATCTACACCTTG    | 1401 |
| HSAJ2425 | NM_017532 | 8923776  | D-003408-02 | GGAAATACGTGGAGACACT   | 1402 |
| HSAJ2425 | NM_017532 | 8923776  | D-003408-03 | CCAGATAACTACGGCGATA   | 1403 |
| HSAJ2425 | NM_017532 | 8923776  | D-003408-04 | TGGCGTACCTTCTCATTGA   | 1404 |
|          |           |          |             |                       |      |
| NR0B1    | NM_000475 | 5016089  | D-003409-01 | CAGCATGGATGATATGATG   | 1405 |
| NR0B1    | NM_000475 | 5016089  | D-003409-02 | CTGCTGAGATTCAATG      | 1406 |
| NR0B1    | NM_000475 | 5016089  | D-003409-03 | ACAGATTCAATCGAACTAA   | 1407 |
| NR0B1    | NM_000475 | 5016089  | D-003409-04 | GAACGTGGCGCTCCTGTAC   | 1408 |
|          |           |          |             |                       |      |
| NR0B2    | NM_021969 | 13259502 | D-003410-01 | GAATATGCCTGCCTGAAAG   | 1409 |
| NR0B2    | NM_021969 | 13259502 | D-003410-02 | GGAATATGCCTGCCTGAAA   | 1410 |
| NR0B2    | NM_021969 | 13259502 | D-003410-03 | CGTAGCCGCTGCCTATGTA   | 1411 |
| NR0B2    | NM_021969 | 13259502 | D-003410-04 | GCCATTCTCTACGCACTTC   | 1412 |
|          |           |          |             |                       |      |
| NR1D1    | NM_021724 | 13430847 | D-003411-01 | CAACACAGGTGGCGTCATCTT | 1413 |
| NR1D1    | NM_021724 | 13430847 | D-003411-02 | GGCATGGTGTACTGTGTATT  | 1414 |
| NR1D1    | NM_021724 | 13430847 | D-003411-03 | CAACATGCATTCGAGAAGTT  | 1415 |
| NR1D1    | NM_021724 | 13430847 | D-003411-04 | GCGCTTGCTTCGTTGTTCTT  | 1416 |
|          |           |          |             |                       |      |
| NR1H2    | NM_007121 | 11321629 | D-003412-01 | GAACAGATCCGGAAGAAGA   | 1417 |
| NR1H2    | NM_007121 | 11321629 | D-003412-02 | GAAGAACAGATCCGGAAGA   | 1418 |
| NR1H2    | NM_007121 | 11321629 | D-003412-03 | CTAAGCAAGTGCCTGGTTT   | 1419 |
| NR1H2    | NM_007121 | 11321629 | D-003412-04 | GCTAACAGCGGCTCAAGAA   | 1420 |

|       |           |          |             |                      |      |
|-------|-----------|----------|-------------|----------------------|------|
|       |           |          |             |                      |      |
| NR1H3 | NM_005693 | 5031892  | D-003413-01 | GAACAGATCCGCCTGAAGA  | 1421 |
| NR1H3 | NM_005693 | 5031892  | D-003413-02 | GGAGATAGTTGACTTGCT   | 1422 |
| NR1H3 | NM_005693 | 5031892  | D-003413-03 | GAGTTGCCTTGCTCATTG   | 1423 |
| NR1H3 | NM_005693 | 5031892  | D-003413-04 | TGACTTGCTAACAGCTA    | 1424 |
| NR1H4 | NM_005123 | 4826979  | D-003414-01 | CAAGTGACCTCGACAACAA  | 1425 |
| NR1H4 | NM_005123 | 4826979  | D-003414-02 | GAAAGAATTGCAAATAGTG  | 1426 |
| NR1H4 | NM_005123 | 4826979  | D-003414-03 | CAACAGACTCTTCTACATT  | 1427 |
| NR1H4 | NM_005123 | 4826979  | D-003414-04 | GAACCATACTCGCAATACA  | 1428 |
| NR1I2 | NM_003889 | 11863133 | D-003415-01 | GAACCATGCTGACTTGTA   | 1429 |
| NR1I2 | NM_003889 | 11863133 | D-003415-02 | GATGGACGCTCAGATGAAA  | 1430 |
| NR1I2 | NM_003889 | 11863133 | D-003415-03 | CAACCTACATGTTCAAAGG  | 1431 |
| NR1I2 | NM_003889 | 11863133 | D-003415-04 | CAGGAGCAATTGCCATTA   | 1432 |
| NR1I3 | NM_005122 | 4826660  | D-003416-01 | GGAAATCTGTCACATCGTA  | 1433 |
| NR1I3 | NM_005122 | 4826660  | D-003416-02 | TCGCAGACATCAACACTTT  | 1434 |
| NR1I3 | NM_005122 | 4826660  | D-003416-03 | CCTCTCGCTACACAATTG   | 1435 |
| NR1I3 | NM_005122 | 4826660  | D-003416-04 | GAACAGTTGTGCAGTTA    | 1436 |
| NR2C1 | NM_003297 | 4507672  | D-003417-01 | TGACAGCACTTGATCATAA  | 1437 |
| NR2C1 | NM_003297 | 4507672  | D-003417-02 | GGAAGGAAGTGTACACCTA  | 1438 |
| NR2C1 | NM_003297 | 4507672  | D-003417-03 | GAGCACATCTCAAACACTAC | 1439 |
| NR2C1 | NM_003297 | 4507672  | D-003417-04 | GAAGAAATTGCACATCAA   | 1440 |
| NR2C2 | NM_003298 | 4507674  | D-003418-01 | GAACAAACGGTGACACTTCA | 1441 |
| NR2C2 | NM_003298 | 4507674  | D-003418-02 | CTGATGAGCTCCAACATAA  | 1442 |
| NR2C2 | NM_003298 | 4507674  | D-003418-03 | CAACCTAACGTGAATCTTG  | 1443 |
| NR2C2 | NM_003298 | 4507674  | D-003418-04 | GAAGACACCTACCGATTGG  | 1444 |
| NR2E1 | NM_003269 | 21361108 | D-003419-01 | GATCATATCTGAAATACAG  | 1445 |
| NR2E1 | NM_003269 | 21361108 | D-003419-02 | CAAGACTGCTTTCAGATAT  | 1446 |
| NR2E1 | NM_003269 | 21361108 | D-003419-03 | GTTAGATGCTACTGAATT   | 1447 |
| NR2E1 | NM_003269 | 21361108 | D-003419-04 | CAATGTATCTCTATGAAGT  | 1448 |
| NR2E3 | NM_014249 | 7657394  | D-003420-01 | GAGAAGCTCCTTGTGATA   | 1449 |
| NR2E3 | NM_014249 | 7657394  | D-003420-02 | GAAGCACTATGGCATCTAT  | 1450 |
| NR2E3 | NM_014249 | 7657394  | D-003420-03 | GAAGGATCCTGAGCACGTA  | 1451 |
| NR2E3 | NM_014249 | 7657394  | D-003420-04 | GAAGCTCCTTGTGATATG   | 1452 |
| NR2F1 | NM_005654 | 20127484 | D-003421-01 | GAAACTCTCATCCGCGATA  | 1453 |
| NR2F1 | NM_005654 | 20127484 | D-003421-02 | TCTCATCCGCGATATGTTA  | 1454 |
| NR2F1 | NM_005654 | 20127484 | D-003421-03 | CAAGAAGTGCCTCAAAGTG  | 1455 |
| NR2F1 | NM_005654 | 20127484 | D-003421-04 | GGAACTTAACCTACACATG  | 1456 |
| NR2F2 | NM_021005 | 14149745 | D-003422-01 | GTACCTGTCCGGATATATT  | 1457 |
| NR2F2 | NM_021005 | 14149745 | D-003422-02 | CCAACCAGCCGACGAGATT  | 1458 |
| NR2F2 | NM_021005 | 14149745 | D-003422-03 | ACTCGTACCTGTCCGGATA  | 1459 |
| NR2F2 | NM_021005 | 14149745 | D-003422-04 | GGCCGTATATGGCAATTCA  | 1460 |
| NR2F6 | NM_005234 | 20070198 | D-003423-01 | CGACGCCGTGGCCTCTCA   | 1461 |
| NR2F6 | NM_005234 | 20070198 | D-003423-02 | CAGCCGGTGTCCGAACTGA  | 1462 |
| NR2F6 | NM_005234 | 20070198 | D-003423-03 | CAACCGTGACTGCCAGATC  | 1463 |
| NR2F6 | NM_005234 | 20070198 | D-003423-04 | GTACTGCCGTCTCAAGAAAG | 1464 |

|       |           |          |             |                      |      |
|-------|-----------|----------|-------------|----------------------|------|
| NR3C1 | NM_000176 | 4504132  | D-003424-01 | GAGGACAGATGTACCACTA  | 1465 |
| NR3C1 | NM_000176 | 4504132  | D-003424-02 | GATAAGACCATGAGTATTG  | 1466 |
| NR3C1 | NM_000176 | 4504132  | D-003424-03 | GAAGACGATTCAATTCTTT  | 1467 |
| NR3C1 | NM_000176 | 4504132  | D-003424-04 | GGACAGATGTACCACTATG  | 1468 |
|       |           |          |             |                      |      |
| NR3C2 | NM_000901 | 4505198  | D-003425-01 | GCAAACAGATGATCCAAGT  | 1469 |
| NR3C2 | NM_000901 | 4505198  | D-003425-02 | CAGCTAACGATTATCAGAA  | 1470 |
| NR3C2 | NM_000901 | 4505198  | D-003425-03 | GCACGAAAGTCAAAGAAGT  | 1471 |
| NR3C2 | NM_000901 | 4505198  | D-003425-04 | GGTATCCGGTCTAGAATA   | 1472 |
|       |           |          |             |                      |      |
| NR4A1 | NM_002135 | 21361341 | D-003426-01 | GAAGGAAAGTTGTCCGAACA | 1473 |
| NR4A1 | NM_002135 | 21361341 | D-003426-02 | CAGGAGAGTTGACACCTT   | 1474 |
| NR4A1 | NM_002135 | 21361341 | D-003426-03 | CAGTGGCTTGACTACTAT   | 1475 |
| NR4A1 | NM_002135 | 21361341 | D-003426-04 | GAAGGCCGCTGTGCTGTG   | 1476 |
|       |           |          |             |                      |      |
| NR4A2 | NM_006186 | 5453821  | D-003427-01 | GCAATGCGTTCGTGGCTTT  | 1477 |
| NR4A2 | NM_006186 | 5453821  | D-003427-02 | CGGCTACACAGGAGAGTTT  | 1478 |
| NR4A2 | NM_006186 | 5453821  | D-003427-03 | CCACGTGACTTTCAACAAT  | 1479 |
| NR4A2 | NM_006186 | 5453821  | D-003427-04 | GAATACAGCTCCGATTCT   | 1480 |
|       |           |          |             |                      |      |
| NR4A3 | NM_006981 | 11276070 | D-003428-01 | CAAAGAAGATCAGACATTA  | 1481 |
| NR4A3 | NM_006981 | 11276070 | D-003428-02 | GATCAGACATTACTTATTG  | 1482 |
| NR4A3 | NM_006981 | 11276070 | D-003428-03 | CCAGAGATCTGATTATTTC  | 1483 |
| NR4A3 | NM_006981 | 11276070 | D-003428-04 | GAAGTTGTCCGTACAGATA  | 1484 |
|       |           |          |             |                      |      |
| NR5A1 | NM_004959 | 20070192 | D-003429-01 | GATTTGAAGTTCCCTGAATA | 1485 |
| NR5A1 | NM_004959 | 20070192 | D-003429-02 | GGAGCGAGCTGCTGGTGTT  | 1486 |
| NR5A1 | NM_004959 | 20070192 | D-003429-03 | GGAGGTGGCCGACCAGATG  | 1487 |
| NR5A1 | NM_004959 | 20070192 | D-003429-04 | CAACGTGCCTGAGCTCATC  | 1488 |
|       |           |          |             |                      |      |
| NR5A2 | NM_003822 | 20070161 | D-003430-01 | CCAAACATATGGCCACTTT  | 1489 |
| NR5A2 | NM_003822 | 20070161 | D-003430-02 | TCAGAGAACTTAAGGTTGA  | 1490 |
| NR5A2 | NM_003822 | 20070161 | D-003430-03 | GGATCCATCTTCCTGGTTA  | 1491 |
| NR5A2 | NM_003822 | 20070161 | D-003430-04 | AAGAATACCTCTACTACAA  | 1492 |
|       |           |          |             |                      |      |
| NR6A1 | NM_033334 | 15451847 | D-003431-01 | CAACGAACCTGTCTCATT   | 1493 |
| NR6A1 | NM_033334 | 15451847 | D-003431-02 | GAAGAACTACACAGATT    | 1494 |
| NR6A1 | NM_033334 | 15451847 | D-003431-03 | GAAGATGGATACGCTGTGA  | 1495 |
| NR6A1 | NM_033334 | 15451847 | D-003431-04 | AAACGATACTGGTACATT   | 1496 |
|       |           |          |             |                      |      |
| null  | D16815    | 2116671  | D-003432-01 | GAAGAATGATCGAATAGAT  | 1497 |
| null  | D16815    | 2116671  | D-003432-02 | GAACATGGAGCAATATAAT  | 1498 |
| null  | D16815    | 2116671  | D-003432-03 | GAGGAGCTTGGCCTTTA    | 1499 |
| null  | D16815    | 2116671  | D-003432-04 | TAAACAAACATGCACTCTGA | 1500 |
|       |           |          |             |                      |      |
| PGR   | NM_000926 | 4505766  | D-003433-01 | GAGATGAGGTCAAGCTACA  | 1501 |
| PGR   | NM_000926 | 4505766  | D-003433-02 | CAGCGTTCTATCAACTTA   | 1502 |
| PGR   | NM_000926 | 4505766  | D-003433-03 | AGATAACTCTCATTAGCTA  | 1503 |
| PGR   | NM_000926 | 4505766  | D-003433-04 | GTAGTCAAGTGGTCAAAT   | 1504 |
|       |           |          |             |                      |      |
| PPARA | NM_005036 | 7549810  | D-003434-01 | TCACGGAGCTCACGGAATT  | 1505 |
| PPARA | NM_005036 | 7549810  | D-003434-02 | GAACATGACATAGAAGATT  | 1506 |
| PPARA | NM_005036 | 7549810  | D-003434-03 | GGATAGTTCTGGAAGCTTT  | 1507 |
| PPARA | NM_005036 | 7549810  | D-003434-04 | GACTCAAGCTGGTATGA    | 1508 |
|       |           |          |             |                      |      |
| PPARD | NM_006238 | 5453939  | D-003435-01 | GAGCGCAGCTGCAAGATTC  | 1509 |

|       |           |          |             |                      |      |
|-------|-----------|----------|-------------|----------------------|------|
| PPARD | NM_006238 | 5453939  | D-003435-02 | GCATGAAGCTGGAGTACGA  | 1510 |
| PPARD | NM_006238 | 5453939  | D-003435-03 | GGAAGCAGTTGGTGAATGG  | 1511 |
| PPARD | NM_006238 | 5453939  | D-003435-04 | GCTGCAAGATTCAAAGAA   | 1512 |
| PPARG | NM_138712 | 20336234 | D-003436-01 | AGACTCAGCTCTACAATAA  | 1513 |
| PPARG | NM_138712 | 20336234 | D-003436-02 | GATTGAAGCTTATCTATGA  | 1514 |
| PPARG | NM_138712 | 20336234 | D-003436-03 | AAGTAACCTCCTCAAATA   | 1515 |
| PPARG | NM_138712 | 20336234 | D-003436-04 | GCATTCTACTCCACATTA   | 1516 |
| RARA  | NM_000964 | 4506418  | D-003437-01 | GACAAGAACTGCATCATCA  | 1517 |
| RARA  | NM_000964 | 4506418  | D-003437-02 | GCAAATACACTACGAACAA  | 1518 |
| RARA  | NM_000964 | 4506418  | D-003437-03 | GAACAACAGCTCAGAACAA  | 1519 |
| RARA  | NM_000964 | 4506418  | D-003437-04 | GAGCAGCAGTCTGAAGAG   | 1520 |
| RARB  | NM_000965 | 14916493 | D-003438-01 | GCACACTGCTCAATCAATT  | 1521 |
| RARB  | NM_000965 | 14916493 | D-003438-02 | GCAGAAGTATTCAAAGAGAA | 1522 |
| RARB  | NM_000965 | 14916493 | D-003438-03 | GGAATGACAGGAACAAGAA  | 1523 |
| RARB  | NM_000965 | 14916493 | D-003438-04 | GCACAGTCCTAGCATCTCA  | 1524 |
| RARG  | NM_000966 | 21359851 | D-003439-01 | GAAATGACCGGAACAAGAA  | 1525 |
| RARG  | NM_000966 | 21359851 | D-003439-02 | TAGAAGAGCTCATCACCAA  | 1526 |
| RARG  | NM_000966 | 21359851 | D-003439-03 | CAAGGAAGCTGTGCGAAAT  | 1527 |
| RARG  | NM_000966 | 21359851 | D-003439-04 | TCAGTGAGCTGGCTACCAA  | 1528 |
| RORA  | NM_134261 | 19743902 | D-003440-01 | GGAAAGAGTTATGTTCTA   | 1529 |
| RORA  | NM_134261 | 19743902 | D-003440-02 | CAAGATCTGTGGAGACAAA  | 1530 |
| RORA  | NM_134261 | 19743902 | D-003440-03 | GCACCTGACTGAAGATGAA  | 1531 |
| RORA  | NM_134261 | 19743902 | D-003440-04 | CCGAGAAGATGGAATACTA  | 1532 |
| RORB  | NM_006914 | 19743906 | D-003441-01 | GCACAGAACATCATTAAAGT | 1533 |
| RORB  | NM_006914 | 19743906 | D-003441-02 | CCACACCTATGAAGAAATT  | 1534 |
| RORB  | NM_006914 | 19743906 | D-003441-03 | GATCAAATTCTACTTCTGA  | 1535 |
| RORB  | NM_006914 | 19743906 | D-003441-04 | TCAAACAGATAAGCAAGA   | 1536 |
| RORC  | NM_005060 | 19743908 | D-003442-01 | TAGAACAGCTGCAGTACAA  | 1537 |
| RORC  | NM_005060 | 19743908 | D-003442-02 | TCACCGAGGCCATTCAAGTA | 1538 |
| RORC  | NM_005060 | 19743908 | D-003442-03 | GAACAGCTGCAGTACAATC  | 1539 |
| RORC  | NM_005060 | 19743908 | D-003442-04 | CCTCATGCCACCTTGAATA  | 1540 |
| RXRA  | NM_002957 | 21536318 | D-003443-01 | TGACGGAGCTTGTGTCCAA  | 1541 |
| RXRA  | NM_002957 | 21536318 | D-003443-02 | CAACAAGGACTGCCTGATT  | 1542 |
| RXRA  | NM_002957 | 21536318 | D-003443-03 | GCAAGGACCTGACCTACAC  | 1543 |
| RXRA  | NM_002957 | 21536318 | D-003443-04 | GCAAGGACCGGAACGAGAA  | 1544 |
| RXRB  | NM_021976 | 21687229 | D-003444-01 | GCAAAGACCTTACATACTC  | 1545 |
| RXRB  | NM_021976 | 21687229 | D-003444-02 | GCAATCATTCTGTAAATC   | 1546 |
| RXRB  | NM_021976 | 21687229 | D-003444-03 | TCACACCGATCCATTGATG  | 1547 |
| RXRB  | NM_021976 | 21687229 | D-003444-04 | GCAAACGGCTATGTGCAAT  | 1548 |
| RXRG  | NM_006917 | 21361386 | D-003445-01 | GGAAGGACCTCATCTACAC  | 1549 |
| RXRG  | NM_006917 | 21361386 | D-003445-02 | CCGGATCTCTGGTAAACA   | 1550 |
| RXRG  | NM_006917 | 21361386 | D-003445-03 | GCGAGCCATTGTACTCTTT  | 1551 |
| RXRG  | NM_006917 | 21361386 | D-003445-04 | GAGCCATTGTACTCTTAA   | 1552 |
| THRA  | NM_003250 | 20127451 | D-003446-01 | GGACAAAGACGAGCAGTGT  | 1553 |
| THRA  | NM_003250 | 20127451 | D-003446-02 | GGAAACAGAGGCAGGAAATT | 1554 |

|      |           |          |             |                      |      |
|------|-----------|----------|-------------|----------------------|------|
| THRA | NM_003250 | 20127451 | D-003446-03 | GTAAGCTGATTGAGCAGAA  | 1555 |
| THRA | NM_003250 | 20127451 | D-003446-04 | GAACCTCCATCCCACCTAT  | 1556 |
| THRΒ | NM_000461 | 10835122 | D-003447-01 | GAATGTCGCTTAAGAAAT   | 1557 |
| THRΒ | NM_000461 | 10835122 | D-003447-02 | GAACAGTCGTCGCCACATC  | 1558 |
| THRΒ | NM_000461 | 10835122 | D-003447-03 | GGACAAGCACCAATAGTCA  | 1559 |
| THRΒ | NM_000461 | 10835122 | D-003447-04 | GTGGAAAGGTTGACTTGGA  | 1560 |
| VDR  | NM_000376 | 4507882  | D-003448-01 | TGAAGAAGCTGAACTTGCA  | 1561 |
| VDR  | NM_000376 | 4507882  | D-003448-02 | GCAACCAAGACTACAAGTA  | 1562 |
| VDR  | NM_000376 | 4507882  | D-003448-03 | TCAATGCTATGACCTGTGA  | 1563 |
| VDR  | NM_000376 | 4507882  | D-003448-04 | CCATTGAGGTACATCATGTT | 1564 |

**Table X**

| Gene Symbol | Sense                | SEQ. ID NO. |
|-------------|----------------------|-------------|
| ABCB1       | GACCAUAAAUGUAAGGUUU  | 1565        |
|             | UAGAAGAUCUGAUGUAAA   | 1566        |
|             | GAAAUGUUCACUUUCAGUUA | 1567        |
|             | GAAGAUCGCUACUGAAGCA  | 1568        |
| ABCC1       | GGAAGCAACUGCAGAGACA  | 1569        |
|             | GAUGACACCUCUCAACAAA  | 1570        |
|             | UAAAGUUGCUCAUCAAGUU  | 1571        |
|             | CAACGAGUCUGCCGAAGGA  | 1572        |
| ABCG2       | GCGAGUGCCUUCUUCGUUA  | 1573        |
|             | AGGCAAUCUUCGUUUAUU   | 1574        |
|             | GGGAAGAAAUCUGGGUCUAA | 1575        |
|             | UGACUCAUCCCAACAUUUA  | 1576        |
| KCNH2       | CCGACGUGCUGCCUGAGUA  | 1577        |
|             | GAGAAGAGCAGCGACACUU  | 1578        |
|             | GAUCAUAGCACCUCUAGAUA | 1579        |
|             | GCUAUUUACUGCUCUUAAU  | 1580        |
|             | UCACUGGGCUCCUUUAUU   | 1581        |
|             | GUGCGAGCCUUCUGAAUUAU | 1582        |
|             | GCUAAGCUAUACUACUGUA  | 1583        |
|             | UGACGGCGCUCUACUUCAC  | 1584        |
| KCNH1       | GAGAUGAAUUCGUUGAAA   | 1585        |
|             | GAAGAACGCAUGAAACGAA  | 1586        |
|             | GAUAAAGACACGAAUUGAAA | 1587        |
|             | GCUGAGAGGUCUAAUAAA   | 1588        |
| CLCA1       | GAACAAACAUGGCUAUGAA  | 1589        |
|             | GUACAUACCUGGCUGGAUU  | 1590        |

|         |                         |      |
|---------|-------------------------|------|
|         | GAACAGCUCACAAGUUAU      | 1591 |
|         | GGAAACGUGUGUCUAUUU      | 1592 |
| SLC6A1  | Sense                   |      |
|         | GGAGGUGGGAGGACAGUUA     | 1593 |
|         | UCACAGCCCUGGUGGAUGA     | 1594 |
|         | GAAGCUGGCUCCUAUGUUC     | 1595 |
|         | GGUACACACUACCAACAUG     | 1596 |
| SLC6A2  | Sense                   |      |
|         | GAACACAAGGUCAACAUUG     | 1597 |
|         | AGAAGGAGCUGGCCUAGUG     | 1598 |
|         | CGGAAACUCUUACAUUUG      | 1599 |
|         | CAACAAAUUUGACAACAAAC    | 1600 |
| SLC21A2 | Sense                   |      |
|         | GUACAUCCUCAUCUUUUU      | 1601 |
|         | GGAAGUGGCUGAGUUUUUA     | 1602 |
|         | GAAGGGAGGCUCAAUGUAA     | 1603 |
|         | GAAGGAAGUGGCUGAGUUA     | 1604 |
| SLC21A3 | Sense                   |      |
|         | GUAGAACAGGAGCUAUUA      | 1605 |
|         | CAAGAUUACUGUCAAACAA     | 1606 |
|         | GCACAAGAGUAUUUGGUAA     | 1607 |
|         | GCAAAUGUCCCCUUCUGUAU    | 1608 |
|         | GCAUGACUCCUAAUAAUAA     | 1609 |
|         | AAACAGCAUUUUCCCUUA      | 1610 |
|         | GAAAAUGCCUCUUCAGGAA     | 1611 |
| SLC28A1 | Sense                   |      |
|         | GUUCAUCGCUCUCCUCUUU     | 1612 |
|         | GGAUCAAGCUGUUUCUGAA     | 1613 |
|         | GGACUGCAGUUUGUACUUG     | 1614 |
|         | GAGUGAACUGACCUAUGG      | 1615 |
| SLC29A1 | Sense                   |      |
|         | GAACGCUGCUCCCGUGGAA     | 1616 |
|         | GAAAGCCACUCUAUCAAAG     | 1617 |
|         | GAAACCAGGUGCCUUCAGA     | 1618 |
|         | CCUCACAGCUGUAUUCAUG     | 1619 |
| SLC26A1 | Sense                   |      |
|         | CCACGGAGCUGCUGGUCAU     | 1620 |
|         | GGGUUGACAUUUUUUGA       | 1621 |
|         | GCACGAGGGUCUCUGUGUU     | 1622 |
|         | GGCCAUCGCCUACUCAUUG     | 1623 |
|         | CAACACCAUGGCAAUUAA      | 1624 |
|         | GAGGAAAGAACUUGCUGAU     | 1625 |
|         | GAGCAAGCGUCCUCCAAAU     | 1626 |
|         | GCAACACCCAUUCAUGGCAAUUA | 1627 |
| SLC26A2 | Sense                   |      |
|         | CCAAAGAACUCAAUGAAC      | 1628 |
|         | ACAAGAACCUUCAGACUAA     | 1629 |
|         | GAAGGUAGAUAGAAGAAUG     | 1630 |

|             |                      |      |
|-------------|----------------------|------|
|             | GUAUUGAACUGUACUGUAA  | 1631 |
| SLC4A4      | Sense                |      |
|             | GCAAUUCUCUUCAUUUAUC  | 1632 |
|             | GGAAAGAUGGUCCACUGAAA | 1633 |
|             | GGACAAAGCCUUCUCAAU   | 1634 |
|             | GGAAUGGGAUCCAGCAUU   | 1635 |
| GLRA1       | Sense                |      |
|             | UGAAAGCCAUGACAUUUG   | 1636 |
|             | CAGACACGGCUGGAGUUUA  | 1637 |
|             | CAAUAGCGCUUCUGGUUU   | 1638 |
|             | GCAGGUAGCAGAUGGACUA  | 1639 |
| KLK1        | Sense                |      |
|             | UCAGAGUGCUGUCUUUAUGU | 1640 |
|             | CAACUUGUUUGACGACGAA  | 1641 |
|             | UGACAGAGCCUGCUGAUAC  | 1642 |
|             | AGGCGGCUCUGUACCAUU   | 1643 |
| ADAM2       | Sense                |      |
|             | GAAACAUGCUGUGAUUUG   | 1644 |
|             | GCAGAUGUUUCCUUUAUA   | 1645 |
|             | CAACAGAGAUGCACUGAUA  | 1646 |
|             | GAAAGGCGCUACAUUGAGA  | 1647 |
| XPNPE<br>P1 | Sense                |      |
|             | GACCUGAGCUUCCCAACAA  | 1648 |
|             | GCGACUGGCUCAACAAUUA  | 1649 |
|             | GAGAUUGCUGGCCUAAUUA  | 1650 |
|             | GACAGCAACUGGACACUUA  | 1651 |
| GZMA        | Sense                |      |
|             | GGAAGAGACUCGUGCAAUG  | 1652 |
|             | GGAACCAUGUGCCAAGUUG  | 1653 |
|             | GAAGUAACUCCUCAUCAA   | 1654 |
|             | GAACUCCUAUAGAUUUCUG  | 1655 |
| CMKLR1      | Sense                |      |
|             | CAUAGAAGCUUACCAAGA   | 1656 |
|             | GAAUGGGAGGAUGAAGAUUA | 1657 |
|             | GGUCAAUGCUCUAAGUGAA  | 1658 |
|             | GAGAGGACUUCUAUGAUG   | 1659 |
| CLN3        | Sense                |      |
|             | CAUC AUGCCUUCUGAAUAA | 1660 |
|             | CAACAGCUCAUCACGAAU   | 1661 |
|             | GCAACAAACUUCUCUUAUGU | 1662 |
|             | GGUCUUCGCUAGCAUCUCA  | 1663 |
| CALCR       | Sense                |      |
|             | GGACCUAGCUGUUGUAAAG  | 1664 |
|             | GAAAGACCAUGCAUUUA    | 1665 |
|             | GCAGGAAGAUGUAUGCUU   | 1666 |
|             | GAUAUAACCAGUAUCGUUA  | 1667 |

|       |                      |      |
|-------|----------------------|------|
| OXTR  | Sense                |      |
|       | GGACCCAGAUAUCCAAUAU  | 1668 |
|       | GCAAUACUAUCCUAACUGA  | 1669 |
|       | GAUUAUAGAUUAGCGUUUG  | 1670 |
|       | GAUGAGGCAUGACUACUAA  | 1671 |
| EDG4  | Sense                |      |
|       | GCGAGUCUGUCCACUUAAC  | 1672 |
|       | GAGAACGCCACCCACUGA   | 1673 |
|       | GAACGGCCACCCACUGAUG  | 1674 |
|       | GGUCAAUGCUGCUGUGUAC  | 1675 |
| EDG5  | Sense                |      |
|       | UCCAGGAACACUAUAUUA   | 1676 |
|       | GUGACCAUCUUCUCCAUCA  | 1677 |
|       | CAUCCUCUGUUGCGCCAUU  | 1678 |
|       | CCAACAAGGUCCAGGAACA  | 1679 |
| EDG7  | Sense                |      |
|       | ACACUGAUACUGUCGAUGA  | 1680 |
|       | AAUAGGAGCAACACUGAUA  | 1681 |
|       | CAGCAGGAGUUACCUUGUU  | 1682 |
|       | GGACACCCAUGAAGCUAAU  | 1683 |
| PTCH  | Sense                |      |
|       | GCACAGAACUCCACUAAA   | 1684 |
|       | GGACAGCAGUUCAUUGUUA  | 1685 |
|       | GAGAAGAGGCUAUGUUUAA  | 1686 |
|       | GGACAAACUUCGACCCUUU  | 1687 |
| SMO   | Sense                |      |
|       | UCGCUACCCUGCUGUUUAU  | 1688 |
|       | GCUACAAGAACUACCGAUA  | 1689 |
|       | CAAGAAAGCUUCCUUCAAC  | 1690 |
|       | GAGAAGAAAUACAGUCAAU  | 1691 |
| CASP3 | Sense                |      |
|       | CAAUUAUCUGAAGAGCUA   | 1692 |
|       | GAACUGGGACUGUGGCAUUG | 1693 |
|       | GUGAGAAGAUGGUUAUUU   | 1694 |
|       | GAGGGUACUUUAAGACAU   | 1695 |
| CASP6 | Sense                |      |
|       | CAUGAGGUGUCAACUGUUA  | 1696 |
|       | GAAGUGAAAUGCUCUAAAUG | 1697 |
|       | AAAUAUUGGCUCCUCCUUAG | 1698 |
|       | GCAAUCACAUUAUAGCAUA  | 1699 |
|       | CAACAUACUGAGGUGGAU   | 1700 |
| CASP7 | Sense                |      |
|       | GAACUCUACUUUAGUCAAU  | 1704 |
|       | GGGCAAUAUGCAUUAUUA   | 1703 |
|       | CAACAGAGGGAGUUUAUUA  | 1704 |
|       | GAACAAAGCCACUGACUGA  | 1705 |

|        |                      |      |
|--------|----------------------|------|
| CASP8  | Sense                |      |
|        | GAAGUGAACUAUGAAGUAA  | 1706 |
|        | CAACAAGGAUGACAAGAAA  | 1707 |
|        | GGACAAAGUUUACCAAAUG  | 1708 |
|        | GAGGGUCGAUCAUCUAUUA  | 1709 |
|        | GAUUAUAGAGGGCUUAUGA  | 1710 |
|        | CAACGACUAUGAAGAAUUC  | 1711 |
|        | GAAGUGAGCAGAUCAGAAU  | 1712 |
| CASP9  | Sense                |      |
|        | CCAGGCAGCUGAUCAUAGA  | 1714 |
|        | UCUCAGGUGUUGCCAAUUA  | 1715 |
|        | GAACAGCUGUAUCAUAGA   | 1716 |
|        | CCACUGGUCUGUAGGGAUU  | 1717 |
| DVL1   | Sense                |      |
|        | UCGUAAAGCUGUUGAUAAUC | 1718 |
|        | GAGGAGAUUUUGAUGACA   | 1719 |
|        | GUAAAGCUGUUGAUAAUCGA | 1720 |
| DVL2   | Sense                |      |
|        | AGACGAAGGUGAUUUACCA  | 1722 |
|        | UGUGAGAGCUACCUAGUCA  | 1723 |
|        | GAAGAAAUUUCAGAUGACA  | 1724 |
| PTEN   | Sense                |      |
|        | GUGAAGAUCUUGACCAAUG  | 1726 |
|        | GAUCAGCAUACACAAUUA   | 1727 |
|        | GAAGUAACCUUCUGCAACA  | 1728 |
| PDK1   | Sense                |      |
|        | GUACAAAGCUGGUUAUAUCC | 1730 |
|        | GAAAGACUCCCAGUGUUA   | 1731 |
|        | GGAAGUCCAUCUCAUCGAA  | 1732 |
| PDK2   | Sense                |      |
|        | GUAAAGAGGGAGACUGAAUG | 1734 |
|        | GGUCUGUGAUGGUCCCUAA  | 1735 |
|        | CAAAGAUGCCUACGACAUG  | 1736 |
| PPP2CA | Sense                |      |
|        | UCACACAAGUUUAUGGUUU  | 1738 |
|        | CAACAGCCGUGACCAUUU   | 1739 |
|        | UAACCAAGCUGCAAUCAUG  | 1740 |
| CTNNA1 | Sense                |      |
|        | GAAGAGAGGUCGUUCUAAG  | 1742 |
|        | AAGCAGAUGUGCAUGAUUA  | 1743 |
|        | UCUAAUACUGCAGUGUUU   | 1744 |

|        |                      |      |
|--------|----------------------|------|
|        | GUAAAGGGCCCUCUAUAA   | 1745 |
| CTNNA2 | Sense                |      |
|        | GAAAGAAUAUGCCCAAGUU  | 1746 |
|        | GAAGAAGAAUGCCACAAUG  | 1747 |
|        | GCAGGAAGAUUAUGAUGUG  | 1748 |
|        | AAAGAAAGCCC AUGUACUA | 1749 |
| HSPCA  | Sense                |      |
|        | GGGAAAGAGCUGCAUAAA   | 1750 |
|        | GCUUAGAACUCUUUACUGA  | 1751 |
|        | UAAAAGAGCUUGACCAAUG  | 1752 |
|        | GCAGAUAAUCUCAUGAUUG  | 1753 |
| DCTN2  | Sense                |      |
|        | CAACUCAUGGUCCAAUACUG | 1754 |
|        | GGAAUGAGCCAGAUGUUUA  | 1755 |
|        | GGAGACAGCUGUACGUUGU  | 1756 |
|        | UCCAAGAGCUGACAAACUGA | 1757 |
| CD2    | Sense                |      |
|        | GUAGGGAGAAGCAAAUAAA  | 1758 |
|        | AAGAUGAGCUUUCUCAUGUA | 1759 |
|        | GGACAUCAUCUCAUCAUU   | 1760 |
|        | GACAAGAGCCCACAGAGUA  | 1761 |
| BAD    | Sense                |      |
|        | GUACUUCCCCUCAGGCCUAU | 1762 |
|        | GCUGUGCCUUGACUACGUA  | 1763 |
|        | GUACUUCCCCUCAGGCCUAU | 1764 |
|        | GGUCAGGUGGCCUCGAGAUC | 1765 |
| SMAC   | Sense                |      |
|        | CAGCGUAACUUCAUUCUUC  | 1766 |
|        | UAACUUCAUUCUUCAGGUA  | 1767 |
|        | CAGCUGCUCUUACCCAUUU  | 1768 |
|        | GAUUGAAGCUUUACUGAA   | 1769 |
|        | UAGAAGAGCUCCGUCAGAA  | 1770 |
|        | CCACAU AUGCGUUGAUUGA | 1771 |
|        | GCGCAGGGCUCUCUACCUA  | 1772 |
| MAP3K5 | Sense                |      |
|        | GAACAGCCUUCAAAUCAAA  | 1773 |
|        | GAUGUUUCUCUACUAUGUUA | 1774 |
|        | GCAAAUACUGGAAGGAAUUA | 1775 |
|        | CAGGAAAGCUCGUAAAUAUA | 1776 |
| PVR    | Sense                |      |
|        | CCACACGGCUGACCUCAUA  | 1777 |
|        | CAGCAGAAUUCUCUUUAUA  | 1778 |
|        | GCAGAAUUCUCUUUAUAUA  | 1779 |
|        | GAUCGGGAUUUAUUUCUAU  | 1780 |
| ERBB2  | Sense                |      |
|        | UGUGGGAGCUGAUGACUUU  | 1781 |
|        | UCACAGAGAUCUUGAAAGG  | 1782 |

|        |                       |      |
|--------|-----------------------|------|
| SOS1   | UGGAAGAGAUCACAGGUUA   | 1783 |
|        | GCUCAUCGCUCACAAACCAA  | 1784 |
| BRCA1  | Sense                 |      |
|        | GAGCACCACUUUCUAUGAUU  | 1785 |
|        | CAAAGAAGCUGUUCAAUUAU  | 1786 |
|        | UGAAAGCCCUCCCUUUAUUA  | 1787 |
|        | GAAGUAGCAUGGAGAAGGA   | 1788 |
| CDKN1A | Sense                 |      |
|        | CCAUCACAGCUUCAAAUAU   | 1789 |
|        | GAAGAGAACUUACUAGUG    | 1790 |
|        | GAAGUGGGCUCCAGUAAUUA  | 1791 |
|        | GCAAGAUGCUGAUUCAUUA   | 1792 |
| CDKN1B | Sense                 |      |
|        | GAACAAAGGAGUCAGACAUU  | 1796 |
|        | AAACUAGGCAGGUUGAAUGA  | 1797 |
|        | GAUGGAAACUUCGACUUUGU  | 1798 |
|        | GUAAACAGAUGGCACUUUG   | 1799 |
| SLC2A4 | Sense                 |      |
|        | GGAAUAGGACAUCCUGUUAUA | 1800 |
|        | GGAGAGAAAGAUGUCAAACGU | 1801 |
|        | GAAUGGACAUCCUGUUAUAA  | 1802 |
|        | GUAAACAGCUCGAAUUAAG   | 1803 |
| NOS2A  | Sense                 |      |
|        | CAGAUAGGCUCCGAAGAUG   | 1804 |
|        | AGACUCAGCUCCAGAAUAC   | 1805 |
|        | GAUCGGGUUCUUUCAUCUUC  | 1806 |
|        | CAGGAUCGGUUCUUUCAUC   | 1807 |
| FRAP1  | Sense                 |      |
|        | GAGCAUGCCGUCAAAUAUA   | 1812 |
|        | CAAGAGAACUCAUCAUUAAG  | 1813 |
|        | CCAAAGUGCUGCAGUACUA   | 1814 |
|        | UAAGAAAGCUAACAAACAA   | 1815 |
| FKBP1A | Sense                 |      |
|        | GAAAACAAGCCCUUUAAGUU  | 1816 |
|        | GAUUUACUCUCCAAGUUGA   | 1817 |
|        | CAGCACAAGUGGUAGGUUA   | 1818 |
|        | GUUGAGGACUGAAUUACUC   | 1819 |
|        | GAUGGCAGCUGUUUAAAUG   | 1820 |
|        | GAGUAUCCUUUCAGUGUUA   | 1821 |

|              |                       |      |
|--------------|-----------------------|------|
| TNFRSF<br>1A | Sense                 |      |
|              | CAAAGGAACCUACUUGUAC   | 1822 |
|              | GGAACCUACUUGUACAAUG   | 1823 |
|              | GAACCUACUUGUACAAUGA   | 1824 |
|              | GAGUGUGUCUCCUGUAGUA   | 1825 |
| IL1R1        | Sense                 |      |
|              | GGACAAGAAUCAAUGGAUA   | 1826 |
|              | GAACAAGCCUCCAGGAUUC   | 1827 |
|              | GGACUUGUGUGGCCUUUAUA  | 1828 |
|              | GAACACAAAGGCACUUAUA   | 1829 |
| IRAK1        | Sense                 |      |
|              | CGAAGAAAGUGAUGAAUUUU  | 1830 |
|              | GCUCUUUGCCCCAUCUCUUUU | 1831 |
|              | UGAAAGACCUGGUGGAAGA   | 1832 |
|              | GCAAUUCAGUUUCUACAUCA  | 1833 |
| TRAF2        | Sense                 |      |
|              | GAAGACAGAGUUAUUAAC    | 1834 |
|              | UCACGAAGACAGAGUUAUU   | 1835 |
|              | AGACAGAGUUAUUAACCA    | 1836 |
|              | CACGAAGACAGAGUUAUUA   | 1837 |
| TRAF6        | Sense                 |      |
|              | CAAAUGAUCUGAGGCAGUU   | 1839 |
|              | GUUCAUAGUUUUGAGCGUUA  | 1840 |
|              | GGAGAAACCUGUUGUGAUU   | 1841 |
|              | GGACAAAGUUGCUGAAUC    | 1842 |
|              | CAAAUGAUCUGAGGCAGUU   | 1843 |
|              | GGAGAAACCUGUUGUGAUU   | 1844 |
|              | GGACAAAGUUGCUGAAUC    | 1845 |
| TRADD        | Sense                 |      |
|              | UGAAGCACCUUGAUCUUUG   | 1847 |
|              | GGGCAGCGCAUACCUGUUU   | 1848 |
|              | GAGGAGCGCUGUUUGAGUU   | 1849 |
|              | GGACGAGGGAGCGCUGUUUG  | 1850 |
|              | GAGGAGCGCUGUUUGAGUU   | 1851 |
|              | GGAUGUCUCUCUCCUCUUU   | 1852 |
|              | GCUCACUCCUUUCUACUAA   | 1853 |
| FADD         | Sense                 |      |
|              | GCACAGAUUUUCCAUUUC    | 1855 |
|              | GCAGUCCUCUUAUUCCUAA   | 1856 |
|              | GAACUCAAGCUGCGUUUAU   | 1857 |
|              | GGACGAAUUGAGAUAAUUAU  | 1858 |
| IKBKE        | Sense                 |      |
|              | UAAGAACACUGCUAUGAA    | 1859 |
|              | GAGGCAUCCUGAAGCAUUA   | 1860 |
|              | GAAGGGCGCUGCAGAACUG   | 1861 |

|             |                      |      |
|-------------|----------------------|------|
|             | GGAACAAGGAGAUCAUGUA  | 1862 |
| IKBKG       | Sense                |      |
|             | CUAUCGAGGUCGUUAAAUU  | 1863 |
|             | GAAUGCAGCUGGAAGAUCU  | 1864 |
|             | GCAGCGAGCUGGACUGUUU  | 1865 |
|             | CCAGACCGAUGUGUAUUUA  | 1866 |
| TNFRSF<br>5 | Sense                |      |
|             | GGUCUCACCUUCGCUAUGGU | 1867 |
|             | GAAAGCGAAUUCGUAGACA  | 1868 |
|             | GCACAAACAAGACUGAUGU  | 1869 |
|             | GAAGGGCACCUCAGAAACA  | 1870 |
|             | UCUCCCAACUUGUAUUAAA  | 1871 |
| RELA        | Sense                |      |
|             | UCAAGUGUCUUCUCAUCAUG | 1872 |
|             | UCAAGUGCCUUAAAUGUAG  | 1873 |
|             | GGAGUACCCUGAGGCUAUA  | 1874 |
|             | GAUGAGAUCUUCUACUGU   | 1875 |
| ARHA        | Sense                |      |
|             | GAGCUGGGCUAAGUAAAUA  | 1876 |
|             | GACCAAAGAUGGAGUGAGA  | 1877 |
|             | GGAAGAAACUGGUGAUUGU  | 1878 |
|             | GGCUGUAACUACUUUAUAA  | 1879 |
| CDC42       | Sense                |      |
|             | GGACAUUUGUUUGGCCAUUU | 1880 |
|             | GGAGAACCAUUAUCUCUUG  | 1881 |
|             | GAACCAAUGCUUUCUCAUG  | 1882 |
|             | GAAGACCUGUUAUGUAGAG  | 1883 |
|             | GAUCAAGAAUUGCAAAUUC  | 1884 |
|             | GGAAAGGGUGACCUAGUA   | 1885 |
|             | UGACAAACCUUUAUGGAAAA | 1886 |
| ROCK1       | Sense                |      |
|             | GGAAUGAGCUUCAGAUGCA  | 1887 |
|             | GGACACAGCUGUAAGAUUG  | 1888 |
|             | GACAAGAGAUUACAGAUAA  | 1889 |
|             | GAAGAAACAUUCCCCUAUUC | 1890 |
| PAK1        | Sense                |      |
|             | GAGGGUGGUUAUGAUUAA   | 1891 |
|             | CAACAAAGAACAAUCACUA  | 1892 |
|             | GAAGAAAUAUACACGGUUU  | 1893 |
|             | UACAUAGACUUUACAGAUA  | 1894 |
| PAK2        | Sense                |      |
|             | GGUAGGAGAUGAAUUGUUU  | 1895 |
|             | AGAAGGAACUGAUCAUUAA  | 1896 |
|             | CUACAGACCUCCAAUAUCA  | 1897 |
|             | GAAACUGGCCAAACCGUUA  | 1898 |
| PAK3        | Sense                |      |

|        |                       |      |
|--------|-----------------------|------|
| PAK4   | Sense                 |      |
|        | GAUUAUCGCUGCAAAGGAA   | 1899 |
|        | GAGAGUGGCCUGCAAGCUU   | 1900 |
|        | GACAAGAGGUGGCCAUAAA   | 1901 |
|        | UUAAAUCGCUGUCUUGAGA   | 1902 |
| PAK6   | Sense                 |      |
|        | ACUAAGAGGUGAACAUAGUA  | 1903 |
|        | GAUCAUGAAUGUCCGAAGA   | 1904 |
|        | GAUGAGACCCUACUACUGA   | 1905 |
|        | CAGCAAAGGUGCCAAAGAU   | 1906 |
| PAK7   | Sense                 |      |
|        | UAAAGGCAGUUGUCCACUA   | 1907 |
|        | GAAGGGACCUGCUUUCUUG   | 1908 |
|        | GCAAAGACGUCCCCUAAGAG  | 1909 |
|        | CCAAUGGGCUGGCUGCAAA   | 1910 |
| HDAC1  | Sense                 |      |
|        | GAGCACGGCUUUAAAAGU    | 1911 |
|        | CAAACUCCGUUAUGAUAAA   | 1912 |
|        | GGAUAAAUGUUGUCUGAUUU  | 1913 |
|        | GGAAAUGGCCUCCAUAAAUA  | 1914 |
| HDAC2  | Sense                 |      |
|        | GGACAUCGCUGUGAAUUGG   | 1915 |
|        | AGAAAAGAACGUACCCGAAGA | 1916 |
|        | GGACAAAGGCCACCCAAUGA  | 1917 |
|        | CCACAGCGAUGACUACAUU   | 1918 |
| CREBBP | Sense                 |      |
|        | GCUGUUAAAUAUGGCUUA    | 1919 |
|        | GCAAAGAAAGCUAGAAUUG   | 1920 |
|        | CAUCAGAGAGUCUUUAUUA   | 1921 |
|        | CCAAUGAGUUGCCAUAAAUA  | 1922 |
| BTRC   | Sense                 |      |
|        | CACAUAAAACUCGUACUUUA  | 1927 |
|        | GAGAAGGCACUCAAGUUUA   | 1928 |
|        | AGACAUAGUUUACAGAGAA   | 1929 |
|        | GCAGAGAGAUUUCAUACU    | 1930 |
| RIPK2  | Sense                 |      |
|        | GAACAUACCUAAAUCAU     | 1931 |
|        | GGACAUCGACCUGUUUAUUA  | 1932 |
|        | UAAAUGAACUCCUACAUAG   | 1933 |
|        | GGAAUUAUCUCUGAACAUUA  | 1934 |
| VAV1   | Sense                 |      |
|        | GCAGAAAUAACAUACUAA    | 1935 |
|        | GCUAUGAGCUGUUCUCAA    | 1936 |

|       |                      |      |
|-------|----------------------|------|
|       | CGACAAAGCUCUACUCAUC  | 1937 |
|       | GCUCAACCCUGGAGACAUU  | 1938 |
| VAV2  | Sense                |      |
|       | GGACAAGACUCGCAGAUUU  | 1939 |
|       | GCUGAGCGCUUGCAAUAA   | 1940 |
|       | CAAGAAGUCUCACGGGAAA  | 1941 |
|       | UCACAGAGGCCAAGAAUU   | 1942 |
| GRB2  | Sense                |      |
|       | UGGAAGCCAUCGCCAAUUA  | 1942 |
|       | CAUCAGUGCAUGACGUUUA  | 1943 |
|       | UGAAUGAGCUGGUGGAAUA  | 1944 |
|       | UGCCAAACUUACCUAUAA   | 1945 |
| PLCG1 | Sense                |      |
|       | GAGCUGCACUCCAUGAGA   | 1946 |
|       | GAAACCAAGCCAUIUAUGA  | 1947 |
|       | CCAAGGAGCUACUGACAUU  | 1948 |
|       | AGAGAAACAUGGCCAAUA   | 1949 |
| ITGB1 | Sense                |      |
|       | CCACAGACAUUUACAUUAA  | 1950 |
|       | GAAGGGAGUUUGCUAAAUU  | 1951 |
|       | GAACAGAUCUGAUGAAUGA  | 1952 |
|       | CAAGAGAGCUGAAGACUAU  | 1953 |
| ITGA4 | Sense                |      |
|       | GCAUUAUAAUUCAGCAUUG  | 1954 |
|       | CAACUUGACUGCAGUUAUG  | 1955 |
|       | GAACUUAACUUUCCAUUU   | 1956 |
|       | GACAAGACCUGUAGUAAUU  | 1957 |
| STAT1 | Sense                |      |
|       | AGAAAAGAGCUUGACAGUAA | 1958 |
|       | GGAAGUAGUUCACAAAAUA  | 1959 |
|       | UGAAGUAUCUGUAUCCAAA  | 1960 |
|       | GAGCUUCACUCCCUUAGUU  | 1961 |
| KRAS2 | Sense                |      |
|       | UAAGGACUCUGAAGAUGUA  | 1962 |
|       | GACAAAGUGUGUAAUUAUG  | 1963 |
|       | GCUCAGGACUUAGCAAGAA  | 1964 |
|       | GAAACUGAAUACCUAAGAU  | 1965 |
|       | GAAACUGAAUACCUAAGAU  | 1966 |
|       | UAAGGACUCUGAAGAUGUA  | 1967 |
|       | GACAAAGUGUGUAAUUAUG  | 1968 |
|       | GCUCAGGACUUAGCAAGAA  | 1969 |
| HRAS  | Sense                |      |
|       | CCAUCCAGCUGAUCCAGAA  | 1970 |
|       | GAACCCUCCUGAUGAGAGU  | 1971 |
|       | GAGGACAUCCACCAGUACA  | 1972 |
| BRAF  | Sense                |      |
|       | GAUUAGAGACCAAGGAAUU  |      |

|        |                      |      |
|--------|----------------------|------|
|        | CCACUGAUGUGUGUUAAUU  | 1973 |
|        | CAAUAGAACCUUGUCAAUAU | 1974 |
|        | GAAGACAGGAAUCGAAUGA  | 1975 |
| ELK1   | Sense                |      |
|        | GAUGUGAGUAGAAGAGUUUA | 1975 |
|        | GGAAGAAUUUUGUACCAUUU | 1976 |
|        | GAACGACCUUUCUUUCUUU  | 1977 |
|        | GGAGUCAUUCUUCUCCUAUA | 1978 |
| RALGDS | Sense                |      |
|        | GGAGAAGCCUCACCUUUG   | 1979 |
|        | GCAGAAAGGACUCAAGAUU  | 1980 |
|        | GAGAACACAUACUCAUUGA  | 1981 |
|        | GAACUUCUCGUCACUGUAU  | 1982 |
| PRKCA  | Sense                |      |
|        | GGAUUGUUCUUUCUCAUA   | 1983 |
|        | GAAGGGUUCUCGUAGUCA   | 1984 |
|        | GAAGAAGGAUGUGGUGAUU  | 1985 |
|        | GGACUGGGGAUCGAACAACA | 1986 |
| MAP2K4 | Sense                |      |
|        | GGACAGAAGUGGAAAUUAUU | 1987 |
|        | UCAAAGAGGUGAACAUUAA  | 1988 |
|        | GACCAAAUCUCAGUUGUUU  | 1989 |
|        | GGAGAAUGGUGCUGUUUAA  | 1990 |
| MAP2K7 | Sense                |      |
|        | GAAGAGACCAAAGUUAUAAU | 1991 |
|        | GAAGACCGGCCACGUCAUU  | 1992 |
|        | GGAAGAGACCAAAGUUAUAA | 1993 |
|        | GCAUUGAGAUUGACCAGAA  | 1994 |
|        | UGAGAGAACGAGAAAGUUG  | 1995 |
|        | GUGAAACCCUGUCUGCAUU  | 1996 |
|        | GGAUUCUCUCUACAACUA   | 1997 |
|        | ACAACUAGGUGAACACAU   | 1998 |
| MAPK8  | Sense                |      |
|        | UCACAGUCCUGAACGGAUA  | 1999 |
|        | GAUUGGAGAUUCUACAUUC  | 2000 |
|        | GCUAUGGAUGCAAAUCUU   | 2001 |
|        | GAAGCUAAGCCGACCAUUU  | 2002 |
| MAPK9  | Sense                |      |
|        | AAAGAGAGCUUAUCGUGAA  | 2003 |
|        | GAUGAUAGGUAGAAAUAG   | 2004 |
|        | ACAAAGAAGUCAUGGGAUUG | 2005 |
|        | GGAGCUGGAUCAUGAAAGA  | 2006 |
| AI1    | Sense                |      |
|        | GAAAAGGGGAUGAUGGGAUU | 2007 |
|        | CCUAGACGAUCCCAAAUAU  | 2008 |
|        | GAGCCAAACCAGGGAUUUA  | 2009 |
|        | UGAAACGAAUGCUGGAGAA  | 2010 |
|        | UCACUCACCCAGAGAAUA   | 2011 |

|         |                      |      |
|---------|----------------------|------|
|         | CCAAGAAAGCUAUCUCUGA  | 2012 |
|         | AGACUCACCUAGAGCUAAA  | 2013 |
| BBC3    | Sense                |      |
|         | CCUGGAGGGGUCCUGUACAA | 2014 |
|         | GAGCAAUAGAGCCAAACGU  | 2015 |
|         | GGAGGGGUCCUGUACAAUCU | 2016 |
|         | GACUUUCUCUGCACCAUGU  | 2017 |
| BCL2L1  | Sense                |      |
|         | CCAGGGAGCUUGAAAGUUU  | 2018 |
|         | AAAGUGCGAGUUCAGUAAA  | 2019 |
|         | GAGAAUCACUAACCAAGAGA | 2020 |
|         | GAGCCCAUCCCCUAAUUAUA | 2021 |
| BCL2L11 | Sense                |      |
|         | GAGACGAGUUUAACGCUUA  | 2022 |
|         | AAAGCAACCUUCUGAUGUA  | 2023 |
|         | CCGAGAAGGUAGACAAUUG  | 2024 |
|         | GCAAAGCAACCUUCUGAUG  | 2025 |
|         | AGACAGAGCCACAAGGUAA  | 2026 |
|         | GCAAGGAGGUUAGAGAAAU  | 2027 |
|         | CAAGGAGGUUAGAGAAUA   | 2028 |
|         | UCUUACGACUGUUACGUUA  | 2029 |
| BID     | Sense                |      |
|         | GAAGACAUCAUCCGGAAUA  | 2030 |
|         | CAACAGCGUUCCUAGAGAA  | 2031 |
|         | GAAAUGGGAUGGACUGAAC  | 2032 |
|         | ACGAUGAGCUGCAGACUGA  | 2033 |
| BIRC2   | Sense                |      |
|         | GAAAGAAGCCUGCAUUAUA  | 2034 |
|         | GAAAUGGACUCUACAUUGU  | 2035 |
|         | ACAAAUAGCACUUAGGUUA  | 2036 |
|         | GAUACACCUGUGGUUAAA   | 2037 |
| BIRC3   | Sense                |      |
|         | GGAGAUGCCUGCCAUUAAA  | 2038 |
|         | UCAAUGAUUCUUGUGUUAGA | 2039 |
|         | GAAAGAACAUUGUAAGUGU  | 2040 |
|         | GAAGAACAUUGUAAG      | 2041 |
| BIRC4   | Sense                |      |
|         | GUAGAUAGAUGGCCAUUAUG | 2042 |
|         | GAGGAGGGCUAACUGAUUG  | 2043 |
|         | GAGGAACCCUGCCAUGUAU  | 2044 |
|         | GCACGGAUUUACUUUUG    | 2045 |
| BIRC5   | Sense                |      |
|         | GGCGUAAGAUGAUGGAAUUA | 2046 |
|         | GCAAAGGAAACCAACAAUA  | 2047 |
|         | GCACAAAGCCAUUCUAAGU  | 2048 |
|         | CAAAGGAAACCAACAAUA   | 2049 |
| BRCA1   | Sense                |      |

|                     |      |
|---------------------|------|
| CCAUACAGCUUCAUAAAUA | 2050 |
| GAAGAGAACUUACUAGUG  | 2051 |
| GAAGUGGGCUCCAGUAAA  | 2052 |
| GCAAGAUGCUGAUUCAUUA | 2053 |
| CCAUACAGCUUCAUAAAUA | 2054 |

|       |                     |      |
|-------|---------------------|------|
| CARD4 | Sense               |      |
|       | GAAAGUUUAUGUCAAGGAA | 2055 |
|       | GAGCAACACUGGCCAUACA | 2056 |
|       | UAACAGAGAUUUGCCUAAA | 2057 |
|       | GCGAAGAGCUGACCAAAUA | 2058 |

|        |                     |      |
|--------|---------------------|------|
| CASP10 | Sense               |      |
|        | CAAAGGGUUUCUCUGUUUA | 2059 |
|        | GAAUUGACCUCCCUCAGUU | 2060 |
|        | GAAGGCAGCUGGUAUUUC  | 2061 |
|        | GACAUGAUCUUCUUCUGA  | 2062 |
|        | GCACUCUUCUGUUCCCCUA | 2063 |

|       |                      |      |
|-------|----------------------|------|
| CASP2 | Sense                |      |
|       | GUAUAAAACUCUCUUUUGA  | 2064 |
|       | GCAAGGAGAUGUCUGAAUA  | 2065 |
|       | CAACUUCCCCUGAUCUUUAA | 2066 |
|       | GCUCAAAGAUGUAUGUAG   | 2067 |

|        |                      |      |
|--------|----------------------|------|
| CDKN1A | Sense                |      |
|        | GAACAAGGAGUCAGACAUU  | 2068 |
|        | AAACUAGGCAGGUUGAAUGA | 2069 |
|        | GAUGGAAUCUUCGACUUUGU | 2070 |
|        | GUAAAACAGAUGGCACUUUG | 2071 |

|       |                     |      |
|-------|---------------------|------|
| CFLAR | Sense               |      |
|       | GAUGUGUCCUCAUUAAAUU | 2072 |
|       | GAAGAGAGAUACAAGAUGA | 2073 |
|       | GAGCAUACCUGAAGAGAGA | 2074 |
|       | GCUAUGAAGUCCAGAAAUU | 2075 |

|      |                      |      |
|------|----------------------|------|
| CLK2 | Sense                |      |
|      | GUGAAUAUGUGAAAUGUG   | 2076 |
|      | AAAGCAUGCUAGAGUAUGA  | 2077 |
|      | UUAAGAAUGUGGAGGAAGUA | 2078 |
|      | GAUAAACAAGCUGACACAU  | 2079 |

|       |                      |      |
|-------|----------------------|------|
| CLSPN | Sense                |      |
|       | GGACGUAAUUGAUGAAGUA  | 2080 |
|       | GCAGAUGGGGUUCUUAAAUG | 2081 |
|       | CAAUGAGGUUGAGGAAAU   | 2082 |
|       | GGAAAUAUCCUGGAGGAUGA | 2083 |

|             |                      |      |
|-------------|----------------------|------|
| CSNK2A<br>1 | Sense                |      |
|             | GAUCCACGUUUCAAUGAUA  | 2084 |
|             | GCAUUUAGGUGGGAGACUUC | 2085 |
|             | GAUGUACGAAUUAUGUUUG  | 2086 |
|             | UGAAUUAGAUCCACGUUUC  | 2087 |

|        |                       |      |
|--------|-----------------------|------|
| CTNNB1 | Sense                 |      |
|        | GCACAAGAAUGGAUCACAA   | 2088 |
|        | GCUGAAACAUUGCAGUUGUA  | 2089 |
|        | GUACGUACCAUGCAGAAUA   | 2090 |
|        | GAACUUGCAUUGUGAUUGG   | 2091 |
| CXCR4  | Sense                 |      |
|        | GAAGCAUGACGGACAAGUA   | 2092 |
|        | GAACAUUCCAGAGCGUGUA   | 2093 |
|        | GUUCUUAGUUGCUGUAUGU   | 2094 |
|        | CAUCAUGGUUGGCCUUAUC   | 2095 |
| CXCR6  | Sense                 |      |
|        | GGAACAAACUGGCAAAGCA   | 2096 |
|        | GAUCAGAGCAGCAGUGAAA   | 2097 |
|        | GGGCAAAACUGAAUUUAUA   | 2098 |
|        | GAUCUCAGGUUCUCCUUGA   | 2099 |
| DAXX   | Sense                 |      |
|        | CUACAGAUCCAAUGAAA     | 2100 |
|        | GCUACAAGCUGGAGAGAUGA  | 2101 |
|        | GGAAACAGCUAUGUGGAAA   | 2102 |
|        | GGAGUUGGAUCUCUCAGAA   | 2103 |
| GAS41  | Sense                 |      |
|        | GUAGUAAGCUAAACUGAAA   | 2104 |
|        | GACAAUAUGUUCAAGAGAA   | 2105 |
|        | GACAACAUCUCGUACAGCUA  | 2106 |
|        | UAUAUGAUGUGUCCAGUAA   | 2107 |
| GTSE1  | Sense                 |      |
|        | CAAAGAACUCACUUACUG    | 2108 |
|        | GAACAGCCCUCUAAAGUGGUU | 2109 |
|        | GAACAUGGAUGACCCUAAG   | 2110 |
|        | GGGCAAAGCUAAAUCAGU    | 2111 |
| HDAC3  | Sense                 |      |
|        | GGAAAGCGAUGUGGAGAUU   | 2112 |
|        | CCAAGACCGUGGCCUAUUU   | 2113 |
|        | AAAGCGAUGUGGAGAUUUA   | 2114 |
|        | GUGAGGAGCUUCCCUAUAG   | 2115 |
| HDAC5  | Sense                 |      |
|        | GAUUUCCUCUUGUGCGAAGU  | 2116 |
|        | GUUAUUAGCACCUUUAAGA   | 2117 |
|        | GGAGGGAGGCCAUGACUUG   | 2118 |
|        | CAGGAGAGCUAAGAAUGG    | 2119 |
|        | GGAUUAUGGAUUUCAGUUAA  | 2120 |
|        | GGAAGUCGGUGCCUUGGUU   | 2121 |
| HEC    | Sense                 |      |
|        | GCAGAUACUUGCACGGUUU   | 2123 |
|        | GAGUAGAACUAGAAUGUGA   | 2124 |
|        | GCGAAUAAAUCAUAGAAAGA  | 2125 |
|        | GAAGAUGGAUUUAUGCAUA   | 2126 |

|               |                       |      |
|---------------|-----------------------|------|
| HIST1H2<br>AA | Sense                 |      |
|               | GGCAAUGCGUCUCGCGAUA   | 2127 |
|               | GAUCCGCAAUGAUGAGGAA   | 2128 |
|               | GCAAUGCGUCUCGCGAUAA   | 2129 |
|               | GAGGAACUCAAUAAGCUUU   | 2130 |
| LMNB1         | Sense                 |      |
|               | AAUAGAACGUGUGCAAUUA   | 2131 |
|               | CAACUGACCUCAUCUGGAA   | 2132 |
|               | GAAGGAAUCUGAUUUAAU    | 2133 |
|               | GGGAAGGGUUUCUCUUAUUA  | 2134 |
| LMNB2         | Sense                 |      |
|               | GGAGGUUCAUUGAGAAUUG   | 2134 |
|               | GGCAAUAGCUCACCGUUUA   | 2135 |
|               | CAAUACGCCUUAGCUGUGU   | 2136 |
|               | GGAGAUCGCCUACAAGUUC   | 2137 |
| MYB           | Sense                 |      |
|               | GCAGAAACACUCCAAUUUA   | 2138 |
|               | GUAAAUAUCGUGAAUGCAUU  | 2139 |
|               | GCACUGAACUUUUGAGAUA   | 2140 |
|               | GAAGAACAGUCAUUUGAUG   | 2141 |
| MYT1          | Sense                 |      |
|               | GAGGUGAGCUGUUAAAUC    | 2142 |
|               | GCAGGGUGAUUUCUAAUUA   | 2143 |
|               | GGGAGAGAUUUUAAUUG     | 2144 |
|               | CAACUUCUCUCCUGAACUU   | 2145 |
| NFKBIB        | Sense                 |      |
|               | GGACACGGCACUGCACUUG   | 2146 |
|               | GCACUUGGCUGUGAUUCAU   | 2148 |
|               | GAGACGAGGGCGAUGAAUA   | 2149 |
|               | CAUGAACCCUUCUGGAUU    | 2150 |
| NFKBIA        | Sense                 |      |
|               | GAACAUUGGACUUGUAUUAU  | 2151 |
|               | GAUGUGGGGUGAAAAGUUA   | 2152 |
|               | GGACGAGAAAGAUCAUUGA   | 2153 |
|               | AGGACGAGCUGCCCCUAGUA  | 2154 |
| NFKBIE        | Sense                 |      |
|               | GAAGGGAAGUUUCAGUAAC   | 2155 |
|               | GGAAGGGAAGUUUCAGUAA   | 2156 |
|               | GGAAACUGCUGCUGUGUAC   | 2157 |
|               | GAACCAACCACUCAUGGAA   | 2158 |
| NUMA1         | Sense                 |      |
|               | GGGAACAGUUUGAAUUAUA   | 2159 |
|               | GCAGUAGCCUGAAGCAGAA   | 2160 |
|               | CGAGAACGGGAUGCACAGAUA | 2161 |
|               | GCAAGAGGCUGAGAGGAAA   | 2162 |
| NUP153        | Sense                 |      |
|               |                       |      |

|             |                       |      |
|-------------|-----------------------|------|
|             | GAAGACAAUAGAAAGCUAA   | 2163 |
|             | GAUAAAGACUGCUGUUAGA   | 2164 |
|             | GAGGAGAGCUCUAUUAUUA   | 2165 |
|             | GAGGAAGCCUGAUUAAAAGA  | 2166 |
| OPA1        | Sense                 |      |
|             | GAAAGAGCAUGAUGACAU    | 2167 |
|             | GAGGAGAGCUCUAUUAUGU   | 2168 |
|             | GAAACUGAAUGGAAGAAUA   | 2169 |
|             | AAAGAAGGCUGUACCGUUA   | 2170 |
| PARVA       | Sense                 |      |
|             | CUACAUGUCUJUGCUCUUA   | 2171 |
|             | GCUAAGUCCUGUAAGAAUA   | 2172 |
|             | CAAAGGCAAUGUACUGUUU   | 2173 |
|             | GAACAAUGGUGGAUCCAAA   | 2174 |
| PIK3CG      | Sense                 |      |
|             | AAGUUCAGCUUCUCUAAUA   | 2175 |
|             | GAAGAAAUCUCUGAUGGAU   | 2176 |
|             | GAACACCUUUACUCUAAUA   | 2177 |
|             | GCAUGGAGCUGGAGAACUA   | 2178 |
| PRKDC       | Sense                 |      |
|             | GAUGAAAGCUCUAAAGAUG   | 2179 |
|             | GAAAGGAGGUUCUAAACUA   | 2180 |
|             | GGAAGAACCUUUUGCUGUUAU | 2181 |
|             | GCAAAGAGGUGGCAGUUA    | 2182 |
| RASA1       | Sense                 |      |
|             | GGAAGAAGAACCAACAU     | 2183 |
|             | GAACAUACUUUCAGAGCUU   | 2184 |
|             | GAACAAUCUUUGCUGUUA    | 2185 |
|             | UAACAGAACUGCUUCAACA   | 2186 |
| SLC9A1      | Sense                 |      |
|             | GAAGAGAUCCACACACAGU   | 2187 |
|             | UCAAUGAGCUGCUGCACAU   | 2188 |
|             | GAAGAUAGGUUUCCAUGUG   | 2189 |
|             | GAUUUACCCUUCCUCAUCU   | 2190 |
| TEGT        | Sense                 |      |
|             | CUACAGAGCUUCAGUGUGA   | 2191 |
|             | GAACAUUUUGAUCGAAAG    | 2192 |
|             | GAGCAAACCUAGAUAAGGA   | 2193 |
|             | GCAUUGAUCUCUUUCUAGA   | 2194 |
| TERT        | Sense                 |      |
|             | GGAAGACAGUGGUGAACUU   | 2195 |
|             | GCAAAGCAUUGGAAUCAGA   | 2196 |
|             | GAGCUGACGUGGAAGAUGA   | 2197 |
|             | GAACGGGCCUGGAACCAUA   | 2198 |
| TNFRSF<br>6 | Sense                 |      |
|             | GAUACUAACUGCUCUCAGA   | 2199 |

|       |                      |      |
|-------|----------------------|------|
|       | GAAAGAAUGGUGUCAUAUGA | 2200 |
|       | UCAAUAAUGUCCCAUGUAA  | 2201 |
|       | UCAUGAAUCUCCAACCUUA  | 2202 |
|       | GAUGUUGACUUGAGUAAA   | 2203 |
| TOP1  | Sense                |      |
|       | GAAAGGAAAUGACUAAUGA  | 2204 |
|       | GAAGAAGGCUGUUCAGAGA  | 2205 |
|       | GGAAGUAGCUACGUUCUUU  | 2206 |
|       | GGACAUAAUGGAAAGAAG   | 2207 |
| TOP2A | Sense                |      |
|       | GAAAGAGGUCCAUCAGAUUU | 2208 |
|       | CAAACUACAUUGGCAUUUA  | 2209 |
|       | AAACAGACAAUGGAUGGAAU | 2210 |
|       | CGAAAGGAAUGGUUAACUA  | 2211 |
| TOP3A | Sense                |      |
|       | CCAGAAAUCUUCCACAGAA  | 2212 |
|       | GAAACUAUCUGGAUGUGUA  | 2213 |
|       | CCACAAAGAUGGUAUCGUA  | 2214 |
|       | GGAAAUGGCUGUGGUAAACA | 2215 |
| TOP3B | Sense                |      |
|       | GAGACAAGAUGAAGACUGU  | 2216 |
|       | GCACAUUGGGCUGCGUCUUU | 2217 |
|       | CCAGUGCGCUUCAAGAUGA  | 2218 |
|       | GAACAUUCUGCUUUGAGGUU | 2219 |
| WEE1  | Sense                |      |
|       | GGUAUUGCCUUGUGAAUUU  | 2220 |
|       | GCAGAACAAUACGAAUAG   | 2221 |
|       | GUACAUAGCUGUUUGAAAU  | 2222 |
|       | GCUGUAAACUUGUAGCAUU  | 2223 |

In addition, to identifying functional siRNA against gene families or pathways, it is possible to design duplexes against genes known to be involved in 5 specific diseases. For example when dealing with human disorders associated with allergies, it will be beneficial to develop siRNA against a number of genes including but not limited to:

the interleukin 4 receptor gene  
 (SEQ. ID NO. 2224: UAGAGGUGCUCAUUCAUU,  
 10 SEQ. ID NO. 2225: GGUAUAAAGCCUUUCCAAGA,  
 SEQ. ID NO. 2225: ACACACAGCUGGAAGAAAU,  
 SEQ. ID NO. 2226: UAACAGAGCUUCCUUAGGU),

the Beta-arrestin-2

(SEQ. ID NO. 2227: GGAUGAAGGAUGACGACUA,  
SEQ. ID NO. 2228: ACACCAACCUCAUUGAAUU,  
SEQ. ID NO. 2229: CGAACAAAGAUGACCAGGUA,  
SEQ. ID NO. 2230: GAUGAAGGAUGACGACUAU, ),

5

the interferon-gamma receptor 1 gene

(SEQ. ID NO. 2231: CAGCAUGGCUCUCCUCUUU,  
SEQ. ID NO. 2232: GUAAAGAACUAUGGUGUUA,  
SEQ. ID NO. 2233: GAAACUACCUGUUACAUUA,

10 SEQ. ID NO. 2234: GAAGUGAGAUCCAGUAUAA),

the matrix metalloproteinase MMP-9

(SEQ. ID NO. 2235: GGAACCAGCUGUAUUUGUU,  
SEQ. ID NO. 2236: GUUGGAGUGUUUCUAAUAA,  
15 SEQ. ID NO. 2237: GCGCUGGGCUUAGAUCAUU,  
SEQ. ID NO. 2238: GGAGCCAGUUUGCCGGUA),

the Slc11a1 (Nramp1) gene

(SEQ. ID NO. 2239: CCAAUGGCCUGCUGAACAA,  
20 SEQ. ID NO. 2240: GGGCCUGGCUUCCUCAUGA,  
SEQ. ID NO. 2241: GGGCAGAGCUCCACCAUGA,  
SEQ. ID NO. 2242: GCACGGCCAUUGCAUUCAA),

SPINK5

25 (SEQ. ID NO. 2243: CCAACUGCCUGUUCAAUAA,  
SEQ. ID NO. 2244: GGAUACAUUGUGAUGAGUUU,  
SEQ. ID NO. 2245: GGACGAAUGUGCUGAGUAU,  
SEQ. ID NO. 2246: GAGCUUGUCUUAUUUGCUA, ),

30 the CYP1A2 gene

(SEQ. ID NO. 2247: GAAAUGCUGUGUCUUCGUA,  
SEQ. ID NO. 2248: GGACAGCACUUCCUGAGA,  
SEQ. ID NO. 2249: GAAGACACCACCAUUCUGA,  
SEQ. ID NO. 2250: GGCCAGAGCUUGACCUUCA),

thymosin-beta4Y

(SEQ. ID NO. 2251: GGACAGGCCUGCGUUGUUU,

SEQ. ID NO. 2252: GGAAAGAGGAAGCUAUGA,

5 SEQ. ID NO. 2253: GCAAACACGUUGGAUGAGU,

SEQ. ID NO. 2254: GGACUAUGCUGGCCUUUJG,

activin A receptor IB

(SEQ. ID NO. 2255: ACAAGACGCUCCAGGAUCU,

10 SEQ. ID NO. 2254: GCAACAGGAUCGACUUGAG,

SEQ. ID NO. 2255: GAAGCUGCGUCCCAACAUC,

SEQ. ID NO. 2256: GCAUAGGCCUGUAAUCGUA,

SEQ. ID NO. 2257: UCAGAGAGUUCGAGACAAA,

SEQ. ID NO. 2258: UGCGAAAGGUUGUAUGUGA,

15 SEQ. ID NO. 2259: GCAACAGGAUCGACUUGAG,

SEQ. ID NO. 2260: GAAUAGCGUUGUGUGUUAU,

SEQ. ID NO. 2261: UGAAUAGCGUUGUGUGUUUA,

SEQ. ID NO. 2262: GGGaucaguuuguugaaUA,

SEQ. ID NO. 2263: GAGCCUGAAUCAUCGUUUA, ),

20

- 28 -

ADAM33

(SEQ. ID NO. 2264: GGAAGUACCUGGAACUGUA,

SEQ. ID NO. 2265: GGACAGAGGGAACCAUUUA,

SEQ. ID NO. 2266: GGUGAGAGGUAGCUCCUAA,

25 SEQ. ID NO. 2267: AAAGACAGGUGGCCACUGA),

the TAP1 gene

(SEQ. ID NO. 2268: GAAAGAUGAUCAGCUAUUU,

SEQ. ID NO. 2269: CAACAGAACCAACAGACAGGU,

30 SEQ. ID NO. 2270: UGAGAAAUGUUCAGAAUGU,

SEQ. ID NO. 2271: UACCUUCACUCGAAACUUA,

COX-2

(SEQ. ID NO. 2272: GAACGAAAGUAAAGAUGUU,

SEQ. ID NO. 2273: GGACUUAUGGGUAAUGUUA,  
SEQ. ID NO. 2274: UGAAAGGACUUAUGGGUAA,  
SEQ. ID NO. 2275: GAUCAGAGUUCACUUUCUU),

5 ADPRT

(SEQ. ID NO. 2276: GGAAAGAUGUUAAGCAUUU,  
SEQ. ID NO. 2277: CAUGGGAGCUCUUGAAAUA,  
SEQ. ID NO. 2278: GAACAAGGAUGAAGUGAAG,  
SEQ. ID NO. 2279: UGAAGAAGCUCACAGUAAA, ),

10

HDC

(SEQ. ID NO. 2280: CAGCAGACCUUCAGUGUGA,  
SEQ. ID NO. 2281: GGAGAGAGAUGGUGGAUUA,  
SEQ. ID NO. 2282: GUACAGAGCUGGAGAUGAA,  
15 SEQ. ID NO. 2283: GAACGUCCCUUCAGUCUGU),

HnmT

(SEQ. ID NO. 2284: CAAAUUCUCUCCAAAGUUC,  
SEQ. ID NO. 2285: GGAUUAUACUGACUGCUUU,  
20 SEQ. ID NO. 2286: GAGCAGAGCUUGGGAAAGA,  
SEQ. ID NO. 2287: GAUAUGAGAUGUAGCAAAU),

GATA-3

(SEQ. ID NO. 2288: GAACUGCUUUCUUUCGUUU,  
25 SEQ. ID NO. 2289: GCAGUAUCAUGAAGCCUAA,  
SEQ. ID NO. 2290: GAAACUAGGUCUGAUAUUC,  
SEQ. ID NO. 2291: GUACAGCUCCGGACUCUUC),

Gab2

30 (SEQ. ID NO. 2292: GCACAACCAUUCUGAAGUU,  
SEQ. ID NO. 2293: GGACUUAAGAUGCCCAGAUG,  
SEQ. ID NO. 2294: GAAGGUGGAUUCUAGGAAA,  
SEQ. ID NO. 2295: GGACUAGCCCUGCUGUUUA), and

## STAT6

(SEQ. ID NO. 2296: GAUAGAAACUCCUGCUAAU,  
SEQ. ID NO. 2297: GGACAUUUAUUCCAGCUA,  
SEQ. ID NO. 2298: GGACAGAGCUACAGACCUA,

5 SEQ. ID NO. 2299: GGAUGGCUCUCCACAGAU).

In addition, rationally designed siRNA or siRNA pools can be directed against genes involved in anemia, hemophilia or hypercholesterolemia. Such genes would include, but are not be limited to:

## 10 APOA5

(SEQ. ID NO. 2300: GAAAGACAGCCUUGAGCAA,  
SEQ. ID NO. 2301: GGACAGGGAGGCCACCAAA,  
SEQ. ID NO. 2302: GGACGAGGCUUGGGCUUUG,  
SEQ. ID NO. 2303: AGCAAGACCUAACAAUAU),

15

## HMG-CoA reductase

(SEQ. ID NO. 2304: GAAUGAAGCUUUGCCUUU,  
SEQ. ID NO. 2305: GAACACAGUUUAGUGCUUU,  
SEQ. ID NO. 2306: UAUCAGAGCUCUUAUGUU,  
20 SEQ. ID NO. 2307: UGAAGAAUGUCUACAGAU),

## NOS3

(SEQ. ID NO. 2308: UGAAGCACCUGGAGAAUGA,  
SEQ. ID NO. 2309: CGGAACAGCACAAGAGUUA,  
25 SEQ. ID NO. 2310: GGAAGAAGACCUUUAAGA,  
SEQ. ID NO. 2309: GCACAAGAGUUUAAGAUC),

## ARH

(SEQ. ID NO. 2310: CGAUACAGCUUGGCACUUU,  
30 SEQ. ID NO. 2311: GAGAAGCGCUGCCCUGUGA,  
SEQ. ID NO. 2312: GAAUCAUGCUGUUCUCUUU,  
SEQ. ID NO. 2313: GGAGUAACCGGACACCUUA),

## CYP7A1

(SEQ. ID NO. 2314: UAAGGUGACUCGAGUGUUU,  
SEQ. ID NO. 2315: AAACGACACUUUCAUCAAA,  
SEQ. ID NO. 2316: GGACUCAAGUUAAGUAUU,  
SEQ. ID NO. 2317: GUAAUGGACUCAAGUAAA),

5

FANCA

(SEQ. ID NO. 2318: GGACAUCAUCUGCCCACUUC,  
SEQ. ID NO. 2319: AGAGGAAGAUGUUCACUUA,  
SEQ. ID NO. 2320: GAUCGUGGCUCUUCAGGAA,  
10 SEQ. ID NO. 2321: GGACAGAGGCAGAUAAAGAA),

FANCG

(SEQ. ID NO. 2322: GCACUAAGCAGCCUCAUG,  
SEQ. ID NO. 2323: GCAAGCAGGUGCCUACAGA,  
15 SEQ. ID NO. 2324: GGAAUUAGAUGCUCCAUUG,  
SEQ. ID NO. 2325: GGACAUCAUCUGCCAAAGUC),

ALAS

(SEQ. ID NO. 2326: CAAUAUGCCUGGAAACUAU,  
20 SEQ. ID NO. 2327: GGUUAAGACUCACCAGUUC,  
SEQ. ID NO. 2328: CAACAGGACUUUAGGUCA,  
SEQ. ID NO. 2329: GCAUAAGAUUGACAUCAUC),

PIGA

25 (SEQ. ID NO. 2330: GAAAGAGGGCAUAAGGUUA,  
SEQ. ID NO. 2331: GGACUGAUCUUAAACUAU,  
SEQ. ID NO. 2332: UCAAAUGGCUUACUCAUC,  
SEQ. ID NO. 2333: UCUAAGAACUGAUGUCUAA), and

30 factor VIII

(SEQ. ID NO. 2334: GCAAAUAGAUCUCCAUUAC,  
SEQ. ID NO. 2335: CCAGAUAAUGUCGUUCUUUA,  
SEQ. ID NO. 2336: GAAAGGCUGUGCUCUAAA,  
SEQ. ID NO. 2337: GGAGAAACCUGCAUGAAAG,

SEQ. ID NO. 2338: CUUGAAGCCUCCUGAAUUA,  
SEQ. ID NO. 2339: GAGGAAGCAUCCAAAGAUU,  
SEQ. ID NO. 2340: GAUAGGAGAUACAAACUUU).

5 Furthermore, rationally designed siRNA or siRNA pools can be directed against genes involved in disorders of the brain and nervous system. Such genes would include, but are not be limited to:

APBB1

(SEQ. ID NO. 2341: CUACGUAGCUCGUGAUAAAG,  
10 SEQ. ID NO. 2342: GCAGAGAUGGUCCACACGUU,  
SEQ. ID NO. 2343: CAUGAGAUCUGCUCUAAGA,  
SEQ. ID NO. 2344: GGGCACCUCUGCUGUAUUG),

BACE1

15 (SEQ. ID NO. 2345: CCACAGAGCAAGUGAUUUAA,  
SEQ. ID NO. 2346: GCAGAAAGGAGAUCAUUUA,  
SEQ. ID NO. 2347: GUAGCAAGAUCUUUACAUAA,  
SEQ. ID NO. 2348: UGUCAGAGCUUGAUUAGAA),

20 PSEN1

(SEQ. ID NO. 2349: GAGCUGACAUUGAAAUAUG,  
SEQ. ID NO. 2350: GUACAGCUAUUUCUCAUCA,  
SEQ. ID NO. 2351: GAGGUUAGGUGAAGUGGUU,  
SEQ. ID NO. 2352: GAAAGGGAGUCACAAGACA,  
25 SEQ. ID NO. 2353: GAACUGGAGUGGAGUAGGA,  
SEQ. ID NO. 2354: CAGCAGGCAUAUCUCAUUA,  
SEQ. ID NO. 2355: UCAAGUACCUCCCUGAAUG),

PSEN2

30 (SEQ. ID NO. 2356: GCUGGGAAGUGGCUUAAUA,  
SEQ. ID NO. 2357: CAUAUUCCCUGCCCUGAUUA,  
SEQ. ID NO. 2358: GGGAAAGUGCUCAAGACCBA,  
SEQ. ID NO. 2359: CAUAGAAAGUGACGUGUUA),

## MASS1

(SEQ. ID NO. 2360: GGAAGGAGCUGUUAUGAGA,  
SEQ. ID NO. 2361: GAAAGGAGAACGUAAAUUA,  
SEQ. ID NO. 2362: GGAGGAAGGUCAAGAUUUA,  
5 SEQ. ID NO. 2363: GGAAAUAGCUGAGAUAAUG, ),

## ARX

(SEQ. ID NO. 2364: CCAGACGCCUGAUAUUGAA,  
SEQ. ID NO. 2365: CAGCACCAUCUAGACCAA,  
10 SEQ. ID NO. 2366: CGCCUGAUAUUGAAGUAAA,  
SEQ. ID NO. 2367: CAACAUCCACUCUCUUCUG) and

## NNMT

(SEQ. ID NO. 2368: GGGCAGUGCUCCAGUGGUAA,  
15 SEQ. ID NO. 2369: GAAAGAGGCUGGUACACAA,  
SEQ. ID NO. 2370: GUACAGAAGUGAGACAUAA,  
SEQ. ID NO. 2371: GAGGUGAUCUCGCAAAGUU).

In addition, rationally designed siRNA or siRNA pools can be directed against  
20 genes involved in hypertension and related disorders. Such genes would include, but  
are not be limited to:

## angiotensin II type 1 receptor

(SEQ. ID NO. 2372: CAAGAAGCCUGCACCAUGU,  
SEQ. ID NO. 2373: GCACUUCACUACCAAAUGA,  
25 SEQ. ID NO. 2374: GCACUGGUCCCAAGUAGUA,  
SEQ. ID NO. 2375: CCAAAGGGCAGUAAAGUUU,  
SEQ. ID NO. 2376: GCUCAGAGGAGGUGUAAUUU,  
SEQ. ID NO. 2377: GCACUUCACUACCAAAUGA,  
SEQ. ID NO. 2378: AAAGGGCAGUAAAGUUU),

30

## AGTR2

(SEQ. ID NO. 2379: GAACAUCUCUGGCAACAAU,  
SEQ. ID NO. 2380: GGUGAUUAUCUCAAAUUG,  
SEQ. ID NO. 2381: GCAAGCAUCUUAUAUAGUU,

SEQ. ID NO. 2382: GAACCAGUCUUCAACUCA), and other related targets.

### Example XIII: Validation of Multigene Knockout using Rab5 and Eps

Two or more genes having similar, overlapping functions often leads to 5 genetic redundancy. Mutations that knockout only one of, *e.g.*, a pair of such genes (also referred to as homologs) results in little or no phenotype due to the fact that the remaining intact gene is capable of fulfilling the role of the disrupted counterpart. To fully understand the function of such genes in cellular physiology, it is often necessary to knockout or knockdown both homologs simultaneously. Unfortunately, 10 concomitant knockdown of two or more genes is frequently difficult to achieve in higher organisms (*e.g.* mice) thus it is necessary to introduce new technologies dissect gene function. One such approach to knocking down multiple genes simultaneously is by using siRNA. For example, **Figure 11** showed that rationally designed siRNA directed against a number of genes involved in the clathrin-mediated endocytosis 15 pathway resulted in significant levels of protein reduction (*e.g.* >80%). To determine the effects of gene knockdown on clathrin-related endocytosis, internalization assays were performed using epidermal growth factor and transferrin. Specifically, mouse receptor-grade EGF (Collaborative Research Inc.) and iron-saturated human transferrin (Sigma) were iodinated as described previously (Jiang, X., Huang, F., 20 Marusyk, A. & Sorkin, A. (2003) *Mol Biol Cell* 14, 858-70). HeLa cells grown in 12-well dishes were incubated with  $^{125}$ I-EGF (1 ng/ml) or  $^{125}$ I-transferrin (1  $\mu$ g/ml) in binding medium (DMEM, 0.1% bovine serum albumin) at 37°C, and the ratio of internalized and surface radioactivity was determined during 5-min time course to calculate specific internalization rate constant  $k_e$  as described previously (Jiang, X *et al.*). The measurements of the uptakes of radiolabeled transferrin and EGF were performed using short time-course assays to avoid influence of the recycling on the uptake kinetics, and using low ligand concentration to avoid saturation of the clathrin-dependent pathway (for EGF Lund, K. A., Opresko, L. K., Strarbuck, C., Walsh, B. J. & Wiley, H. S. (1990) *J. Biol. Chem.* 265, 15713-13723).

30

The effects of knocking down Rab5a, 5b, 5c, Eps, or Eps 15R (individually) are shown in **Figure 22** and demonstrate that disruption of single genes has little or no effect on EGF or Tfn internalization. In contrast, simultaneous knock down of Rab5a,

5b, and 5c, or Eps and Eps 15R, leads to a distinct phenotype (note: total concentration of siRNA in these experiments remained constant with that in experiments in which a single siRNA was introduced, see **Figure 23**). These experiments demonstrate the effectiveness of using rationally designed siRNA to 5 knockdown multiple genes and validates the utility of these reagents to override genetic redundancy.

**Example XIV. Validation of Multigene Targeting Using G6PD, GAPDH, PLK, and UQC.**

10 Further demonstration of the ability to knock down expression of multiple genes using rationally designed siRNA was performed using pools of siRNA directed against four separate genes. To achieve this, siRNA were transfected into cells (total siRNA concentration of 100nM) and assayed twenty-four hours later by B-DNA. Results shown in **Figure 24** show that pools of rationally designed molecules are 15 capable of simultaneously silencing four different genes.

**Example XV. Validation of Multigene Knockouts As Demonstrated by Gene Expression Profiling, a Prophetic Example**

20 To further demonstrate the ability to concomitantly knockdown the expression of multiple gene targets, single siRNA or siRNA pools directed against a collection of genes (e.g. 4, 8, 16, or 23 different targets) are simultaneously transfected into cells and cultured for twenty-four hours. Subsequently, mRNA is harvested from treated (and untreated) cells and labeled with one of two fluorescent probes dyes (e.g. a red fluorescent probe for the treated cells, a green fluorescent probe for the control cells.). 25 Equivalent amounts of labeled RNA from each sample is then mixed together and hybridized to sequences that have been linked to a solid support (e.g. a slide, “DNA CHIP”). Following hybridization, the slides are washed and analyzed to assess changes in the levels of target genes induced by siRNA.

30 **Example XVI. Identifying Hyperfunctional siRNA**

**Identification of Hyperfunctional Bcl-2 siRNA**

The ten rationally designed Bcl2 siRNA (identified in **Figure 13, 14**) were tested to identify hyperpotent reagents. To accomplish this, each of the ten Bcl-2

siRNA were individually transfected into cells at a 300pM (0.3nM) concentrations. Twenty-four hours later, transcript levels were assessed by B-DNA assays and compared with relevant controls. As shown in **Figure 25**, while the majority of Bcl-2 siRNA failed to induce functional levels of silencing at this concentration, siRNA 1 and 8 induced >80% silencing, and siRNA 6 exhibited greater than 90% silencing at this subnanomolar concentration.

By way of prophetic examples, similar assays could be performed with any of the groups of rationally designed genes described in Example VII or Example VIII.

10 Thus for instance, rationally designed siRNA sequences directed against PDGFA

(SEQ. ID NO. 2383: GGUAAGAUUAUUGUGCUUUA,

SEQ. ID NO. 2384: CCGCAAAUAUGCAGAAUUA,

SEQ. ID NO. 2385: GGAUGUACAUUUGCGUGUUA,

15 SEQ. ID NO. 2386: GGUGAAGUUUGUAUGUUUA), or

PDGFB

(SEQ. ID NO. 2387: GCUCCGCGCUUUCGAAUUU,

SEQ. ID NO. 2388: GAGCAGGAAUUGGUGAGAGAUG,

20 SEQ. ID NO. 2389: GAACUUGGGAUAAAGAGAGU,

SEQ. ID NO. 2390: CCGAGGGAGCUUUAUGAGAGAU,

SEQ. ID NO. 2391: UUUAUGAGAUGCUGAGUGA)

could be introduced into cells at increasingly limiting concentrations to determine whether any of the duplexes are hyperfunctional. Similarly, rationally designed 25 sequences directed against

HIF1 alpha

(SEQ. ID NO. 2392: GAAGGAACCUGAUGCUUUA,

SEQ. ID NO. 2393: GCAUAUAUCUAGAAGGUAU,

SEQ. ID NO. 2394: GAACAAAUAACAUGGGAUUA,

30 SEQ. ID NO. 2395: GGACACAGAUUAGACUUG), or

VEGF

(SEQ. ID NO. 2396: GAACGUACUUGCAGAUGUG,

SEQ. ID NO. 2397: GAGAAAGCAUUUGUUUGUA,

SEQ. ID NO. 2398: GGAGAAAGCAUUUGUUUGU,

SEQ. ID NO. 2399: CGAGGCAGCUUGAGUUAAA) could be introduced into cells at increasingly limiting concentrations and screened for hyperfunctional duplexes.

5 **Example XVII: Gene Silencing: Prophetic Example**

Below is an example of how one might transfet a cell.

- a. Select a cell line. The selection of a cell line is usually determined by the desired application. The most important feature to RNAi is the level of expression of the gene of interest. It is highly recommended to use cell lines for which siRNA transfection conditions have been specified and validated.
- 10 b. Plate the cells. Approximately 24 hours prior to transfection, plate the cells at the appropriate density so that they will be approximately 70 – 90% confluent, or approximately  $1 \times 10^5$  cells/ml at the time of transfection. Cell densities that are too low may lead to toxicity due to excess exposure and uptake of transfection reagent-siRNA complexes. Cell densities that are too high may lead to low transfection efficiencies and little or no silencing. Incubate the cells overnight. Standard incubation conditions for mammalian cells are 37°C in 5% CO<sub>2</sub>. Other cell types, such as insect cells, require different 15 temperatures and CO<sub>2</sub> concentrations that are readily ascertainable by persons skilled in the art. Use conditions appropriate for the cell type of interest.
- 20 c. SiRNA re-suspension. Add 20 µl siRNA universal buffer to each siRNA to generate a final concentration of 50 µM.
- 25 d. SiRNA-lipid complex formation. Use RNase-free solutions and tubes. Using the following table, Table XI:

e.

Table XI

|   | 96-well                       | 24-well                       |
|---|-------------------------------|-------------------------------|
| <b>Mixture 1 (TransIT-TKO-Plasmid dilution mixture)</b>     |                               |                               |
| Opti-MEM  | 9.3 $\mu$ l                   | 46.5 $\mu$ l                  |
| TransIT-TKO (1 $\mu$ g/ $\mu$ l)                            | 0.5 $\mu$ l                   | 2.5 $\mu$ l                   |
| <b>Mixture 1 Final Volume</b>                               | <b>10.0 <math>\mu</math>l</b> | <b>50.0 <math>\mu</math>l</b> |
| <b>Mixture 2 (siRNA dilution mixture)</b>                   |                               |                               |
| Opti-MEM  | 9.0 $\mu$ l                   | 45.0 $\mu$ l                  |
| siRNA (1 $\mu$ M)   | 1.0 $\mu$ l                   | 5.0 $\mu$ l                   |
| <b>Mixture 2 Final Volume</b>                               | <b>10.0 <math>\mu</math>l</b> | <b>50.0 <math>\mu</math>l</b> |
| <b>Mixture 3 (siRNA-Transfection reagent mixture)</b>       |                               |                               |
| Mixture 1   | 10 $\mu$ l                    | 50 $\mu$ l                    |
| Mixture 2   | 10 $\mu$ l                    | 50 $\mu$ l                    |
| <b>Mixture 3 Final Volume</b>                               | <b>20 <math>\mu</math>l</b>   | <b>100 <math>\mu</math>l</b>  |
| <b>Incubate 20 minutes at room temperature.</b>             |                               |                               |
| <b>Mixture 4 (Media-siRNA/Transfection reagent mixture)</b> |                               |                               |
| Mixture 3   | 20 $\mu$ l                    | 100 $\mu$ l                   |
| Complete media  | 80 $\mu$ l                    | 400 $\mu$ l                   |
| <b>Mixture 4 Final Volume</b>                               | <b>100 <math>\mu</math>l</b>  | <b>500 <math>\mu</math>l</b>  |
| <b>Incubate 48 hours at 37°C.</b>                           |                               |                               |

5 **Transfection.** Create a Mixture 1 by combining the specified amounts of OPTI-MEM serum free media and transfection reagent in a sterile polystyrene tube. Create a Mixture 2 by combining specified amounts of each siRNA with OPTI-MEM media in sterile 1 ml tubes. Create a Mixture 3 by combining specified amounts of Mixture 1 and Mixture 2. Mix gently (do not vortex) and incubate at room temperature for 20 minutes. Create a Mixture 4 by combining specified amounts of Mixture 3 to complete media. Add appropriate volume to each cell culture well. Incubate cells with transfection reagent mixture for 24 – 72 hours at 37°C. This incubation time is flexible. The ratio of silencing will remain consistent at any point in the time period. Assay for gene silencing using an appropriate detection method such as RT-PCR, Western blot analysis, immunohistochemistry, phenotypic analysis, mass

10

15

spectrometry, fluorescence, radioactive decay, or any other method that is now known or that comes to be known to persons skilled in the art and that from reading this disclosure would be useful with the present invention. The optimal window for observing a knockdown phenotype is related to the mRNA turnover of the gene of interest, although 24 – 72 hours is standard. Final Volume reflects amount needed in each well for the desired cell culture format. When adjusting volumes for a Stock Mix, an additional 10% should be used to accommodate variability in pipetting, *etc.* Duplicate or triplicate assays should be carried out when possible.

10 While the invention has been described in connection with specific embodiments thereof, it will be understood that it is capable of further modifications and this application is intended to cover any variations, uses, or adaptations of the invention following, in general, the principles of the invention and including such departure from the present disclosure as come within known or customary practice 15 within the art to which the invention pertains and as may be applied to the essential features hereinbefore set forth and as follows in the scope of the appended claims.

**Claims**

1. A method for selecting siRNA comprising selecting an siRNA molecule of 19 – 25 nucleoside bases, said method comprising:

5 (a) selecting a target gene;

(b) measuring the functionality of sequences of 19 – 25 nucleotides in length that are substantially complementary to a stretch of nucleotides of the target sequence, wherein said functionality is dependent upon non-target specific criteria.

10 2. The method according to claim 1 wherein said functionality is determined by applying one of the following formulas:

Formula I = -(GC/3) +(AU<sub>15-19</sub>) -(Tm<sub>20°C</sub>)\*3 -(G<sub>13</sub>)\*3 -(C<sub>19</sub>) +(A<sub>19</sub>)\*2 +(A<sub>3</sub>) +(U<sub>10</sub>)+(A<sub>14</sub>) -(U<sub>5</sub>) -(A<sub>11</sub>);

15 Formula II = -(GC/3) -(AU<sub>15-19</sub>)\*3 -(G<sub>13</sub>)\*3 -(C<sub>19</sub>) +(A<sub>19</sub>)\*2 +(A<sub>3</sub>);

Formula III = -(GC/3) +(AU<sub>15-19</sub>) -(Tm<sub>20°C</sub>)\*3;

20 Formula IV = -(GC/2)+(AU<sub>15-19</sub>)/2-(Tm<sub>20°C</sub>)\*2 -(G<sub>13</sub>)\*3 -(C<sub>19</sub>) +(A<sub>19</sub>)\*2 +(A<sub>3</sub>) +(U<sub>10</sub>)+(A<sub>14</sub>) -(U<sub>5</sub>) -(A<sub>11</sub>);

Formula V = -(G<sub>13</sub>)\*3 -(C<sub>19</sub>)+(A<sub>19</sub>)\*2 +(A<sub>3</sub>) +(U<sub>10</sub>)+(A<sub>14</sub>) -(U<sub>5</sub>) -(A<sub>11</sub>);

25 Formula VI = -(G<sub>13</sub>)\*3 -(C<sub>19</sub>) +(A<sub>19</sub>)\*2 +(A<sub>3</sub>);

Formula VII = -(GC/2)+(AU<sub>15-19</sub>)/2 -(Tm<sub>20°C</sub>)\*1 -(G<sub>13</sub>)\*3 -(C<sub>19</sub>) +(A<sub>19</sub>)\*3 +(A<sub>3</sub>)\*3 +(U<sub>10</sub>)/2+(A<sub>14</sub>)/2 -(U<sub>5</sub>)/2 -(A<sub>11</sub>)/2;

30

wherein in Formulas I – VII:

AU<sub>15-19</sub> = 0 – 5 depending on the number of A or U bases on the sense

strand at positions 15–19;

$G_{13} = 1$  if G is the base at position 13 on the sense strand, otherwise its value is 0;

$C_{19} = 1$  if C is the base at position 19 of the sense strand, otherwise its value is 0;

5

$GC =$  the number of G and C bases in the entire sense strand;

$Tm_{20^\circ C} = 1$  if the  $Tm$  is greater than  $20^\circ C$ ;

$A_3 = 1$  if A is the base at position 3 on the sense strand, otherwise its value is 0;

10

$A_{11} = 1$  if A is the base at position 11 on the sense strand, otherwise its value is 0;

$A_{14} = 1$  if A is the base at position 14 on the sense strand, otherwise its value is 0;

15

$A_{19} = 1$  if A is the base at position 19 on the sense strand, otherwise its value is 0;

$U_5 = 1$  if U is the base at position 5 on the sense strand, otherwise its value is 0;

$U_{10} = 1$  if U is the base at position 10 on the sense strand, otherwise its value is 0;

20

or,

Formula VIII:  $(-14)*G_{13}-13*A_1-12*U_{7-11}*U_{2-10}*A_{11}-10*U_4-10*C_3-10*C_5-10*C_6-$

$9*A_{10}-9*C_{18}-8*G_{10}-7*U_1-7*U_{16}-7*C_{17}-7*C_{19}$

25

$+7*U_{17}+8*A_2+8*A_4+8*A_5+8*C_4+9*G_8+10*A_7+10*U_{18}+11*A_{19}+$   
 $11*C_9+15*G_1+18*A_3+19*U_{10}-Tm-3*(GC_{total})-6*(GC_{15-19})-$   
 $30*X;$  and

30

Formula IX:  $(14.1)*A_3+(14.9)*A_6+(17.6)*A_{13}+(24.7)*A_{19}+(14.2)*U_{10}+(10.5)*$

$C_9+(23.9)*G_1+(16.3)*G_2+(-12.3)*A_{11}+(-19.3)*U_1+(-12.1)*U_2+$

$(-11)*U_3+(-15.2)*U_{15}+(-11.3)*U_{16}+(-11.8)*C_3+(-17.4)*C_6+(-$

$10.5)*C_7+(-13.7)*G_{13}+(-25.9)*G_{19}-Tm-3*(GC_{total})-6*(GC_{15-19})-$

$30*X$

wherein

$A_1 = 1$  if A is the base at position 1 of the sense strand, otherwise its value is 0;  
 $A_2 = 1$  if A is the base at position 2 of the sense strand, otherwise its value is 0;  
 $A_3 = 1$  if A is the base at position 3 of the sense strand, otherwise its value is 0;  
 $A_4 = 1$  if A is the base at position 4 of the sense strand, otherwise its value is 0;  
5       $A_5 = 1$  if A is the base at position 5 of the sense strand, otherwise its value is 0;  
 $A_6 = 1$  if A is the base at position 6 of the sense strand, otherwise its value is 0;  
 $A_7 = 1$  if A is the base at position 7 of the sense strand, otherwise its value is 0;  
 $A_{10} = 1$  if A is the base at position 10 of the sense strand, otherwise its value is 0;  
 $A_{11} = 1$  if A is the base at position 11 of the sense strand, otherwise its value is 0;  
10      $A_{13} = 1$  if A is the base at position 13 of the sense strand, otherwise its value is 0;  
 $A_{19} = 1$  if A is the base at position 19 of the sense strand, otherwise if another base  
is present or the sense strand is only 18 base pairs in length, its value is 0;

$C_3 = 1$  if C is the base at position 3 of the sense strand, otherwise its value is 0;  
15      $C_4 = 1$  if C is the base at position 4 of the sense strand, otherwise its value is 0;  
 $C_5 = 1$  if C is the base at position 5 of the sense strand, otherwise its value is 0;  
 $C_6 = 1$  if C is the base at position 6 of the sense strand, otherwise its value is 0;  
 $C_7 = 1$  if C is the base at position 7 of the sense strand, otherwise its value is 0;  
 $C_9 = 1$  if C is the base at position 9 of the sense strand, otherwise its value is 0;  
20      $C_{17} = 1$  if C is the base at position 17 of the sense strand, otherwise its value is 0;  
 $C_{18} = 1$  if C is the base at position 18 of the sense strand, otherwise its value is 0;  
 $C_{19} = 1$  if C is the base at position 19 of the sense strand, otherwise if another base  
is present or the sense strand is only 18 base pairs in length, its value is 0;

25      $G_1 = 1$  if G is the base at position 1 on the sense strand, otherwise its value is 0;  
 $G_2 = 1$  if G is the base at position 2 of the sense strand, otherwise its value is 0;  
 $G_8 = 1$  if G is the base at position 8 on the sense strand, otherwise its value is 0;  
 $G_{10} = 1$  if G is the base at position 10 on the sense strand, otherwise its value is 0;  
 $G_{13} = 1$  if G is the base at position 13 on the sense strand, otherwise its value is 0;  
30      $G_{19} = 1$  if G is the base at position 19 of the sense strand, otherwise if another base  
is present or the sense strand is only 18 base pairs in length, its value is 0;

$U_1 = 1$  if U is the base at position 1 on the sense strand, otherwise its value is 0;  
 $U_2 = 1$  if U is the base at position 2 on the sense strand, otherwise its value is 0;

$U_3 = 1$  if U is the base at position 3 on the sense strand, otherwise its value is 0;

$U_4 = 1$  if U is the base at position 4 on the sense strand, otherwise its value is 0;

$U_7 = 1$  if U is the base at position 7 on the sense strand, otherwise its value is 0;

$U_9 = 1$  if U is the base at position 9 on the sense strand, otherwise its value is 0;

5       $U_{10} = 1$  if U is the base at position 10 on the sense strand, otherwise its value is 0;

$U_{15} = 1$  if U is the base at position 15 on the sense strand, otherwise its value is 0;

$U_{16} = 1$  if U is the base at position 16 on the sense strand, otherwise its value is 0;

$U_{17} = 1$  if U is the base at position 17 on the sense strand, otherwise its value is 0;

$U_{18} = 1$  if U is the base at position 18 on the sense strand, otherwise its value is 0;

10

$GC_{15-19}$  = the number of G and C bases within positions 15 – 19 of the sense strand or within positions 15 – 18 if the sense strand is only 18 base pairs in length;

$GC_{total}$  = the number of G and C bases in the sense strand;

$Tm = 100$  if the targeting site contains an inverted repeat longer than 4 base pairs, otherwise its value is 0; and

15

$X$  = the number of times that the same nucleotide repeats four or more times in a row.

20

3.      A method of gene-silencing comprising selecting an siRNA according to claim 2 and introducing it into a cell.

25

4.      The method according to claim 3 wherein said introducing is by allowing passive uptake of the siRNA. ~~passive uptake~~

30

5.      The method according to claim 3, wherein said introducing is through the use of a vector.

6.      A method for developing an siRNA algorithm for selecting siRNA, said method comprising:

30

(a) selecting a set of siRNA;

(b) measuring the gene silencing ability of each siRNA from said set;

(c) determining the relative functionality of each siRNA;

(d) determining the amount of improved functionality by the presence or absence of at least one variable selected from the group consisting of

the total GC content, melting temperature of the siRNA, GC content at positions 15 –19, the presence or absence of a particular nucleotide at a particular position and the number of times that the same nucleotide repeats within a given sequence; and

5 (e) developing an algorithm using the information of step (d).

7. A method of selecting an siRNA with improved functionality, said method comprising using the algorithm of claim 6.

10 8. A method of selecting hyperfunctional siRNA, said method comprising using at least one functional siRNA, wherein at least one said functional siRNA has been selected according to the method of claim 7 and measuring the silencing ability of said at least one functional siRNA, wherein silencing ability is measured at a concentration of less than 1 15 nanomolar siRNA.

9. An siRNA molecule, wherein said siRNA molecule is effective at silencing Bcl-2.

20 10. The siRNA molecule of claim 9, wherein said siRNA molecule comprises a sequence substantially similar to a sequence selected from the group consisting of GGGAGAUAGUGAUGAAGUA (SEQ. ID NO. 301); GAAGUACAUCAUUUAAG (SEQ. ID NO. 302); GUACGACAACCGGGAGAUA (SEQ. ID NO. 303); AGAUAGUGAUGAAGUACAU (SEQ. ID NO. 304); UGAAGACUCUGCUCAGUUU (SEQ. ID NO. 305); CAUGCGGCCUCUGUUUGAUUU (SEQ. ID NO. 306); UGCGGCCUCUGUUUGAUUU (SEQ. ID NO. 307); GAGAUAGUGAUGAAGUACA (SEQ. ID NO. 308); 25 GGAGAUAGUGAUGAAGUAC (SEQ. ID NO. 309); and GAAGACUCUGCUCAGUUUG (SEQ. ID NO. 310).

30 11. The siRNA molecule of claim 10, wherein said siRNA molecule comprises a sequence selected from the group consisting of

GGGAGAUAGUGAUGAAGUA (SEQ. ID NO. 301);  
GAAGUACAUCCAUUAUAAG (SEQ. ID NO. 302);  
GUACGACAACCGGGAGAUA (SEQ. ID NO. 303);  
AGAUAGUGAUGAAGUACAU (SEQ. ID NO. 304);  
5 UGAAGACUCUGCUCAGUUU (SEQ. ID NO. 305);  
GCAUGCGGCCUCUGUUUGA (SEQ. ID NO. 306);  
UGCGGCCUCUGUUUGAUUU (SEQ. ID NO. 307);  
GAGAUAGUGAUGAAGUACA (SEQ. ID NO. 308);  
GGAGAUAGUGAUGAAGUAC (SEQ. ID NO. 309); and  
10 GAAGACUCUGCUCAGUUUG (SEQ. ID NO. 310).

12. The siRNA molecule of claim 11, wherein said siRNA molecule comprises GCAUGCGGCCUCUGUUUGA .
15. The siRNA molecule of claim 9, wherein said siRNA molecule comprises a sense strand and an anti-sense strand.
14. The siRNA molecule of claim 9, wherein said siRNA molecule comprises a hairpin.
20. The siRNA molecule of claim 9, wherein said siRNA molecule comprises between 18 and 30 base pairs.
25. A kit for gene silencing comprising at least one siRNA selected from the group consisting of sequences substantially similar to the group consisting of GGGAGAUAGUGAUGAAGUA (SEQ. ID NO. 301);  
GAAGUACAUCCAUUAUAAG (SEQ. ID NO. 302);  
GUACGACAACCGGGAGAUA (SEQ. ID NO. 303);  
AGAUAGUGAUGAAGUACAU (SEQ. ID NO. 304);  
30 UGAAGACUCUGCUCAGUUU (SEQ. ID NO. 305);  
GCAUGCGGCCUCUGUUUGA (SEQ. ID NO. 306);  
UGCGGCCUCUGUUUGAUUU (SEQ. ID NO. 307);  
GAGAUAGUGAUGAAGUACA (SEQ. ID NO. 308);

GGAGAUAGUGAUGAAGUAC (SEQ. ID NO. 309); and  
GAAGACUCUGCUCAGUUUG (SEQ. ID NO. 310).

17. A method of gene silencing comprising using the siRNA molecule of  
5 claim 10.

18. A method of gene silencing comprising using the siRNA molecule of claim  
11.

10 19. A kit, wherein said kit is comprised of at least two siRNA, wherein said at  
least two siRNA comprise a first optimized siRNA and a second optimized  
siRNA, wherein said first optimized siRNA and said second optimized  
siRNA are optimized according to one of the following formulas:

15 Formula I = -(GC/3) +(AU<sub>15-19</sub>) -(Tm<sub>20°C</sub>)\*3 -(G<sub>13</sub>)\*3 -(C<sub>19</sub>) +(A<sub>19</sub>)\*2 +(A<sub>3</sub>)  
+(U<sub>10</sub>)+(A<sub>14</sub>)-(U<sub>5</sub>)-(A<sub>11</sub>);

Formula II = -(GC/3) -(AU<sub>15-19</sub>)\*3 -(G<sub>13</sub>)\*3 -(C<sub>19</sub>) +(A<sub>19</sub>)\*2 +(A<sub>3</sub>);

20 Formula III = -(GC/3) +(AU<sub>15-19</sub>) -(Tm<sub>20°C</sub>)\*3;

Formula IV = -(GC/2)+(AU<sub>15-19</sub>)/2-(Tm<sub>20°C</sub>)\*2 -(G<sub>13</sub>)\*3 -(C<sub>19</sub>) +(A<sub>19</sub>)\*2  
+(A<sub>3</sub>) +(U<sub>10</sub>)+(A<sub>14</sub>)-(U<sub>5</sub>)-(A<sub>11</sub>);

25 Formula V = -(G<sub>13</sub>)\*3 -(C<sub>19</sub>) +(A<sub>19</sub>)\*2 +(A<sub>3</sub>) +(U<sub>10</sub>)+(A<sub>14</sub>)-(U<sub>5</sub>)-(A<sub>11</sub>);

Formula VI = -(G<sub>13</sub>)\*3 -(C<sub>19</sub>) +(A<sub>19</sub>)\*2 +(A<sub>3</sub>);

30 Formula VII = -(GC/2) +(AU<sub>15-19</sub>)/2 -(Tm<sub>20°C</sub>)\*1 -(G<sub>13</sub>)\*3 -(C<sub>19</sub>) +(A<sub>19</sub>)\*3  
+(A<sub>3</sub>)\*3 +(U<sub>10</sub>)/2+(A<sub>14</sub>)/2 -(U<sub>5</sub>)/2 -(A<sub>11</sub>)/2;

wherein in Formulas I – VII:

$AU_{15-19} = 0 - 5$  depending on the number of A or U bases on the sense strand at positions 15 -19;

5  $G_{13} = 1$  if G is the base at position 13 on the sense strand, otherwise its value is 0;

$C_{19} = 1$  if C is the base at position 19 of the sense strand, otherwise its value is 0;

$GC =$  the number of G and C bases in the entire sense strand;

$Tm_{20^{\circ}C} = 1$  if the  $Tm$  is greater than  $20^{\circ}C$ ;

10  $A_3 = 1$  if A is the base at position 3 on the sense strand, otherwise its value is 0;

$A_{11} = 1$  if A is the base at position 11 on the sense strand, otherwise its value is 0;

15  $A_{14} = 1$  if A is the base at position 14 on the sense strand, otherwise its value is 0;

$A_{19} = 1$  if A is the base at position 19 on the sense strand, otherwise its value is 0;

$U_5 = 1$  if U is the base at position 5 on the sense strand, otherwise its value is 0;

20  $U_{10} = 1$  if U is the base at position 10 on the sense strand, otherwise its value is 0;

or,

25 Formula VIII:  $(-14)*G_{13}-13*A_1-12*U_7-11*U_2-10*A_{11}-10*U_4-10*C_3-10*C_5-10*C_6-9*A_{10}-9*U_9-9*C_{18}-8*G_{10}-7*U_1-7*U_{16}-7*C_{17}-7*C_{19}+7*U_{17}+8*A_2+8*A_4+8*A_5+8*C_4+9*G_8+10*A_7+10*U_{18}+11*A_{19}+11*C_9+15*G_1+18*A_3+19*U_{10}-Tm-3*(GC_{total})-6*(GC_{15-19})-30*X;$  and

30

Formula IX:  $(14.1)*A_3+(14.9)*A_6+(17.6)*A_{13}+(24.7)*A_{19}+(14.2)*U_{10}+(10.5)*C_9+(23.9)*G_1+(16.3)*G_2+(-12.3)*A_{11}+(-19.3)*U_1+(-12.1)*U_2+$

$$(-11)*U_3 + (-15.2)*U_{15} + (-11.3)*U_{16} + (-11.8)*C_3 + (-17.4)*C_6 + (-10.5)*C_7 + (-13.7)*G_{13} + (-25.9)*G_{19} - Tm - 3*(GC_{total}) - 6*(GC_{15-19}) - 30*X$$

wherein

5       $A_1 = 1$  if A is the base at position 1 of the sense strand, otherwise its value is 0;  
 $A_2 = 1$  if A is the base at position 2 of the sense strand, otherwise its value is 0;  
 $A_3 = 1$  if A is the base at position 3 of the sense strand, otherwise its value is 0;  
 $A_4 = 1$  if A is the base at position 4 of the sense strand, otherwise its value is 0;  
 $A_5 = 1$  if A is the base at position 5 of the sense strand, otherwise its value is 0;  
10      $A_6 = 1$  if A is the base at position 6 of the sense strand, otherwise its value is 0;  
 $A_7 = 1$  if A is the base at position 7 of the sense strand, otherwise its value is 0;  
 $A_{10} = 1$  if A is the base at position 10 of the sense strand, otherwise its value is 0;  
 $A_{11} = 1$  if A is the base at position 11 of the sense strand, otherwise its value is 0;  
 $A_{13} = 1$  if A is the base at position 13 of the sense strand, otherwise its value is 0;  
15      $A_{19} = 1$  if A is the base at position 19 of the sense strand, otherwise if another base  
          is present or the sense strand is only 18 base pairs in length, its value is 0;

20      $C_3 = 1$  if C is the base at position 3 of the sense strand, otherwise its value is 0;  
 $C_4 = 1$  if C is the base at position 4 of the sense strand, otherwise its value is 0;  
 $C_5 = 1$  if C is the base at position 5 of the sense strand, otherwise its value is 0;  
 $C_6 = 1$  if C is the base at position 6 of the sense strand, otherwise its value is 0;  
 $C_7 = 1$  if C is the base at position 7 of the sense strand, otherwise its value is 0;  
 $C_9 = 1$  if C is the base at position 9 of the sense strand, otherwise its value is 0;  
 $C_{17} = 1$  if C is the base at position 17 of the sense strand, otherwise its value is 0;  
25      $C_{18} = 1$  if C is the base at position 18 of the sense strand, otherwise its value is 0;  
 $C_{19} = 1$  if C is the base at position 19 of the sense strand, otherwise if another base  
          is present or the sense strand is only 18 base pairs in length, its value is 0;

30      $G_1 = 1$  if G is the base at position 1 on the sense strand, otherwise its value is 0;  
 $G_2 = 1$  if G is the base at position 2 of the sense strand, otherwise its value is 0;  
 $G_8 = 1$  if G is the base at position 8 on the sense strand, otherwise its value is 0;  
 $G_{10} = 1$  if G is the base at position 10 on the sense strand, otherwise its value is 0;  
 $G_{13} = 1$  if G is the base at position 13 on the sense strand, otherwise its value is 0;

$G_{19} = 1$  if G is the base at position 19 of the sense strand, otherwise if another base is present or the sense strand is only 18 base pairs in length, its value is 0;

U<sub>1</sub> = 1 if U is the base at position 1 on the sense strand, otherwise its value is 0;  
5 U<sub>2</sub> = 1 if U is the base at position 2 on the sense strand, otherwise its value is 0;  
U<sub>3</sub> = 1 if U is the base at position 3 on the sense strand, otherwise its value is 0;  
U<sub>4</sub> = 1 if U is the base at position 4 on the sense strand, otherwise its value is 0;  
U<sub>7</sub> = 1 if U is the base at position 7 on the sense strand, otherwise its value is 0;  
U<sub>9</sub> = 1 if U is the base at position 9 on the sense strand, otherwise its value is 0;  
10 U<sub>10</sub> = 1 if U is the base at position 10 on the sense strand, otherwise its value is 0;  
U<sub>15</sub> = 1 if U is the base at position 15 on the sense strand, otherwise its value is 0;  
U<sub>16</sub> = 1 if U is the base at position 16 on the sense strand, otherwise its value is 0;  
U<sub>17</sub> = 1 if U is the base at position 17 on the sense strand, otherwise its value is 0;  
U<sub>18</sub> = 1 if U is the base at position 18 on the sense strand, otherwise its value is 0;

15

GC<sub>15-19</sub> = the number of G and C bases within positions 15 – 19 of the sense strand or within positions 15 – 18 if the sense strand is only 18 base pairs in length;

GC<sub>total</sub> = the number of G and C bases in the sense strand;

20 Tm = 100 if the targeting site contains an inverted repeat longer than 4 base pairs, otherwise its value is 0; and

X = the number of times that the same nucleotide repeats four or more times in a row.

Figure 1

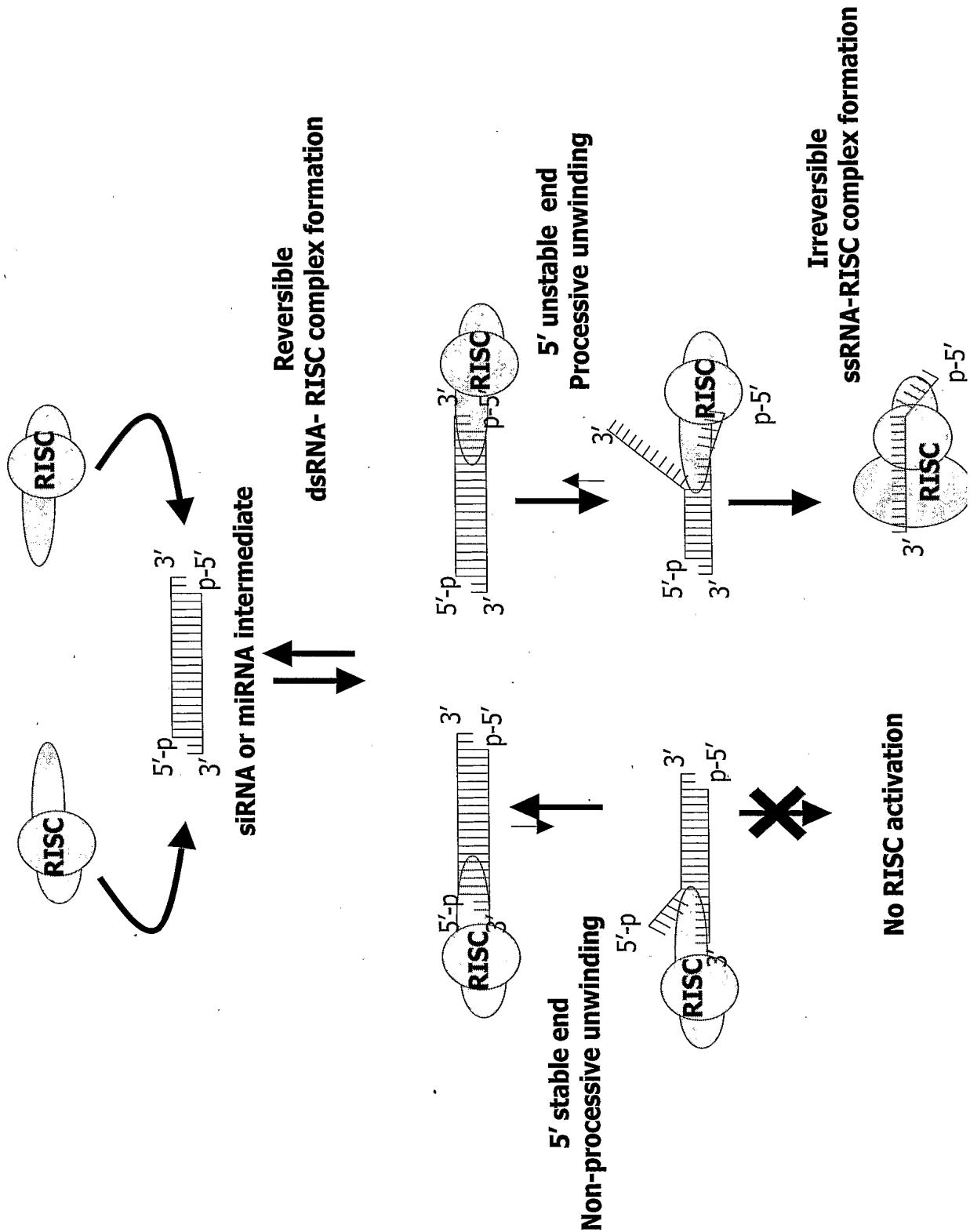
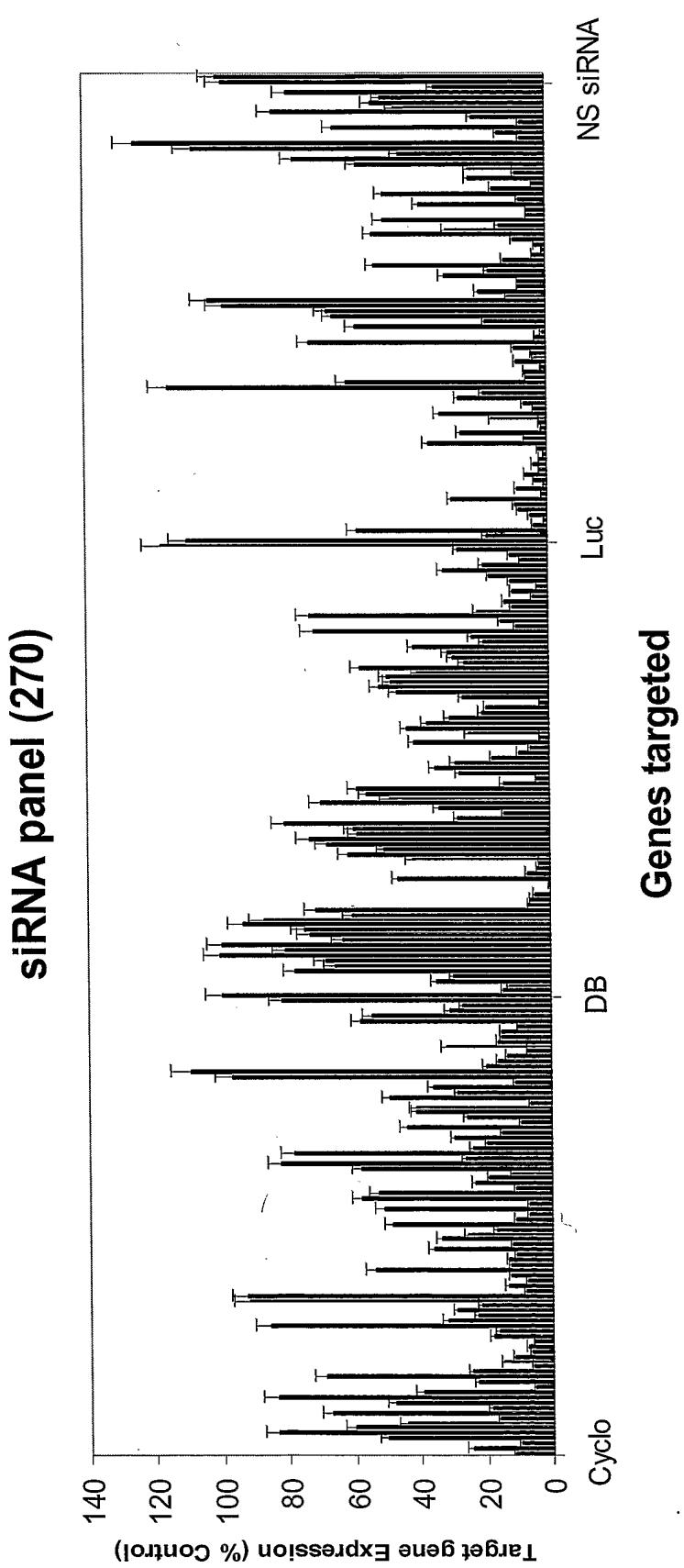
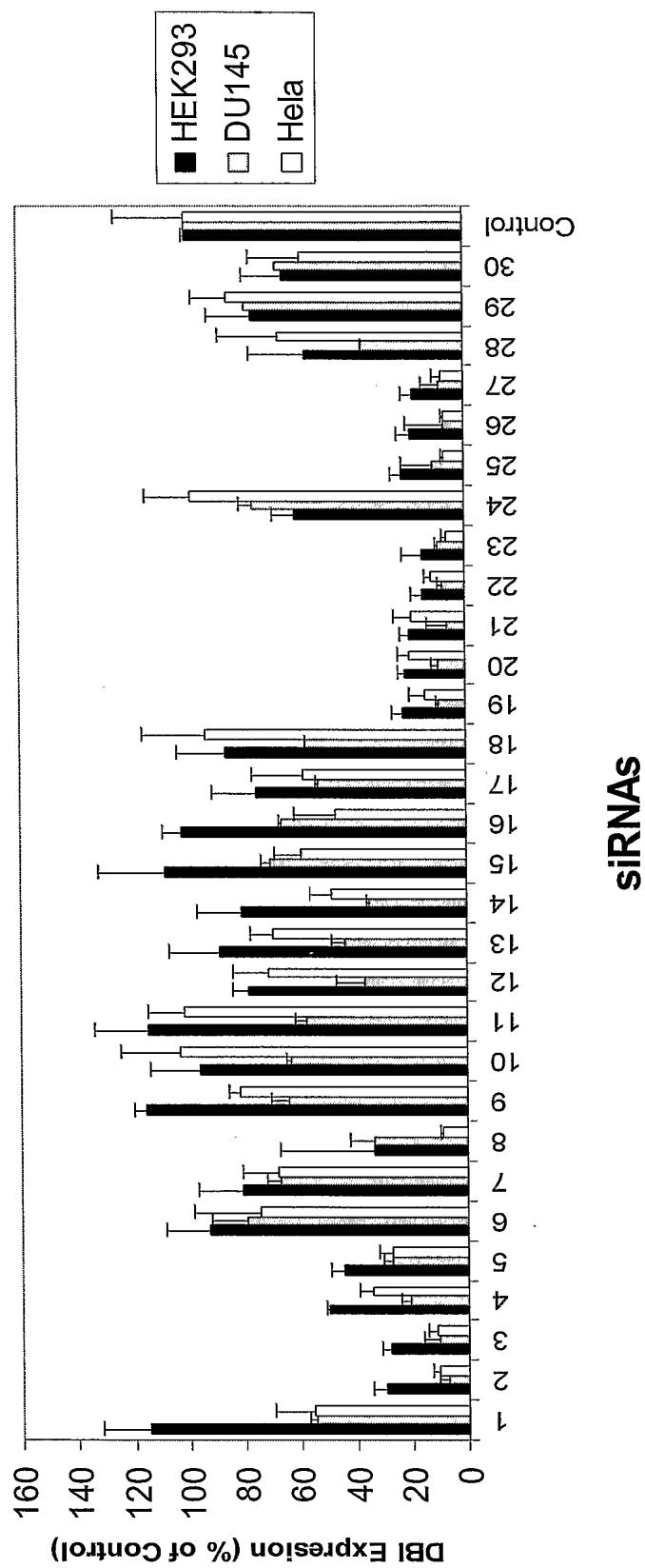


Figure 2



**siRNA functionality is independent from the cell line**

Figure 3a



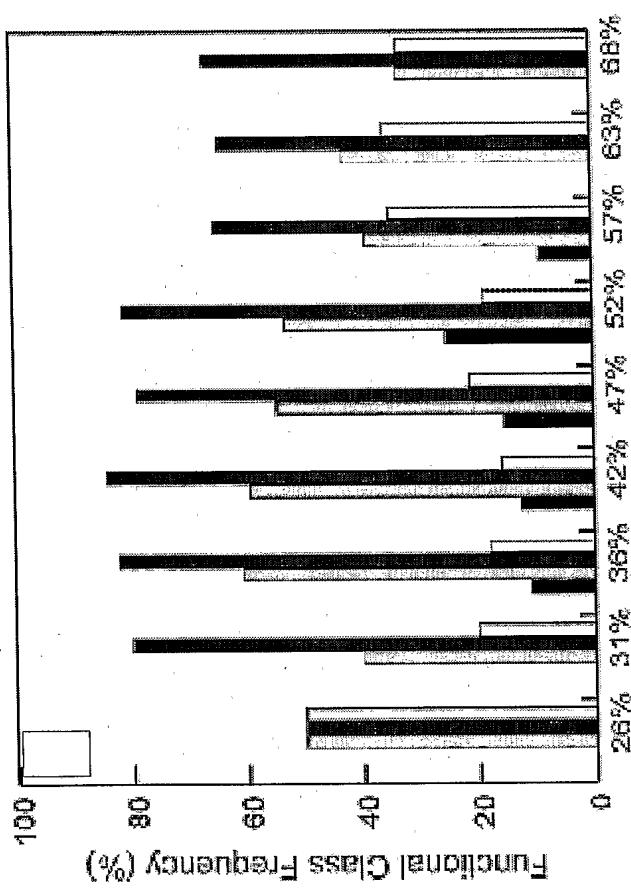


Figure 3b

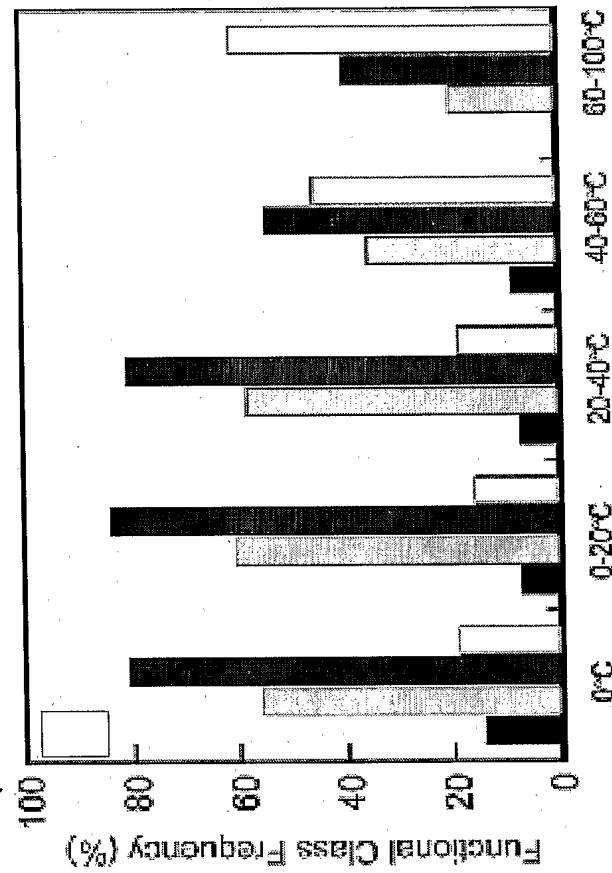


Figure 3c

Figure 4

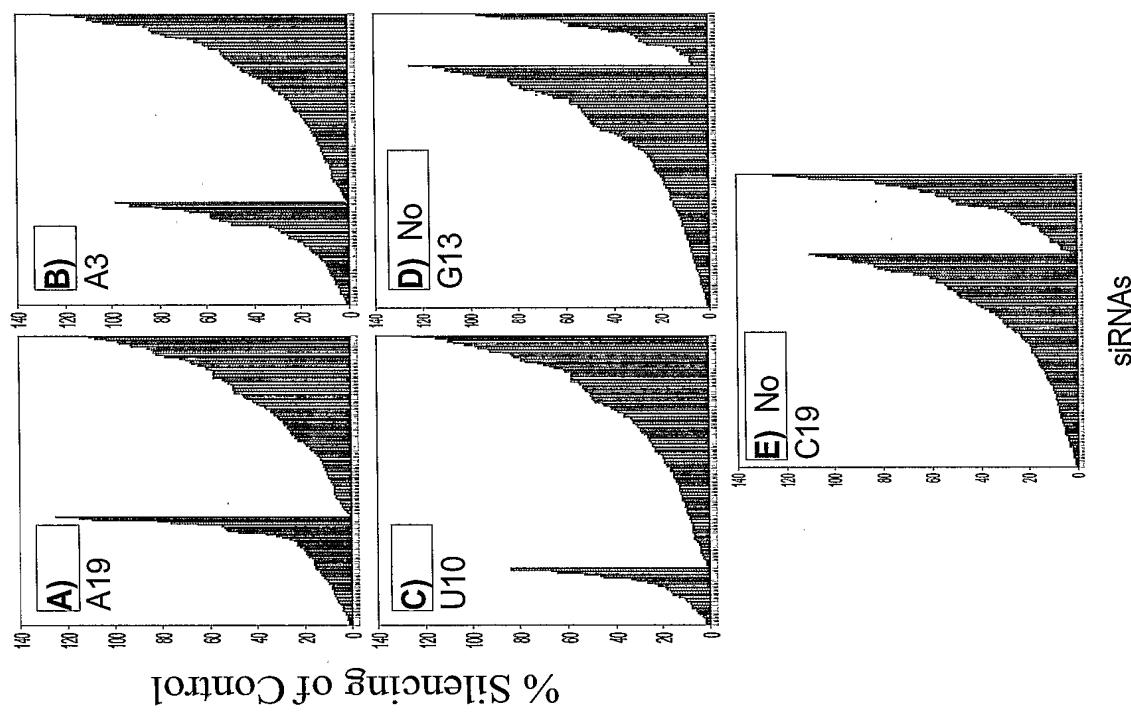


Figure 5A

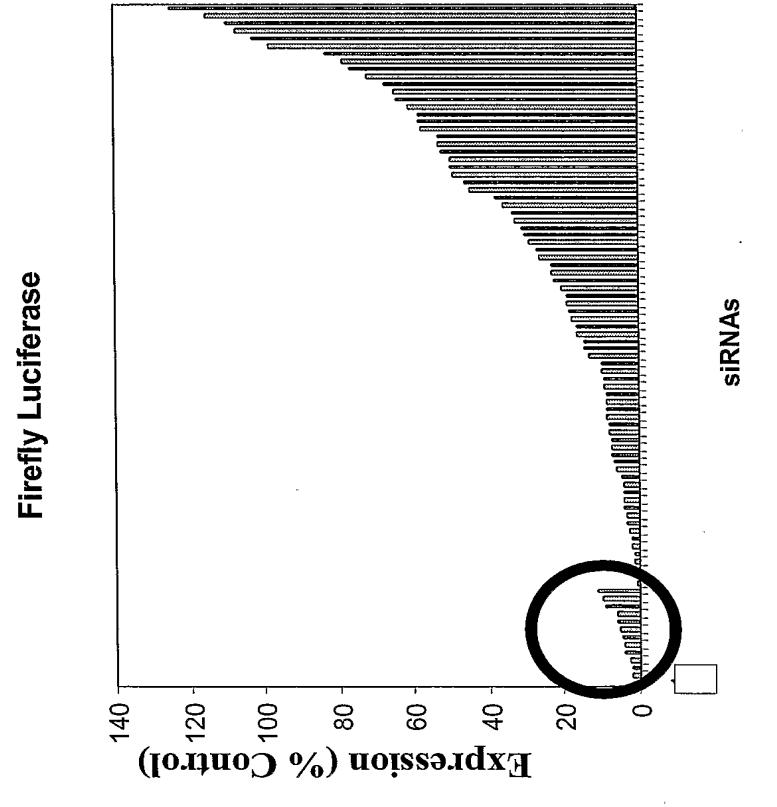
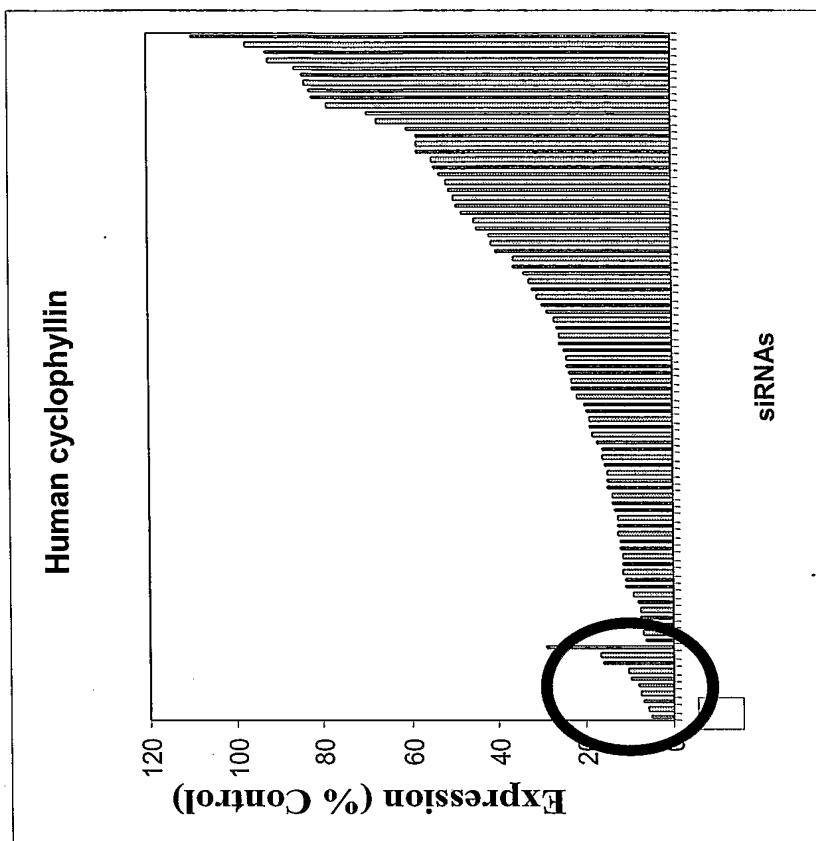


Figure 5B



**Differential internal stability**

Figure 6a

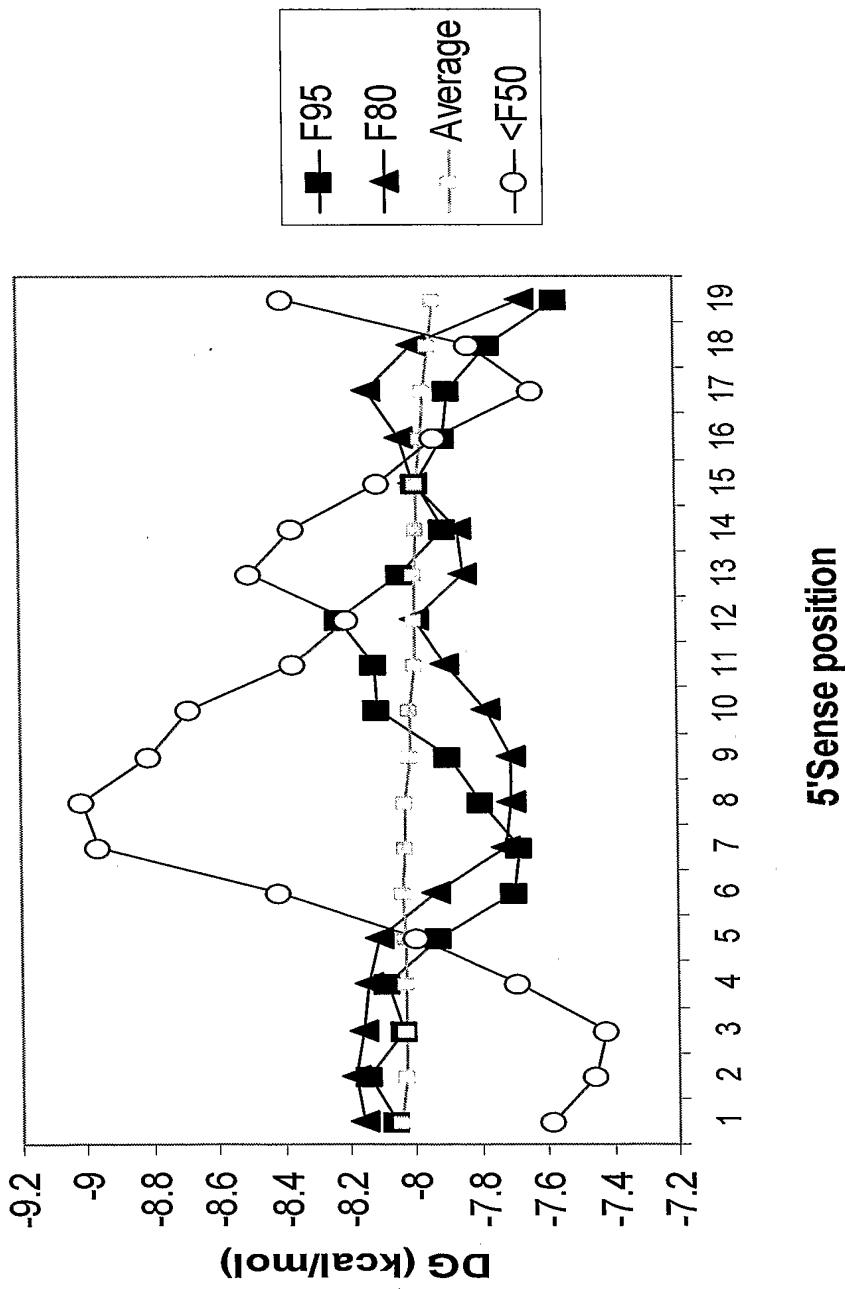


Figure 6b

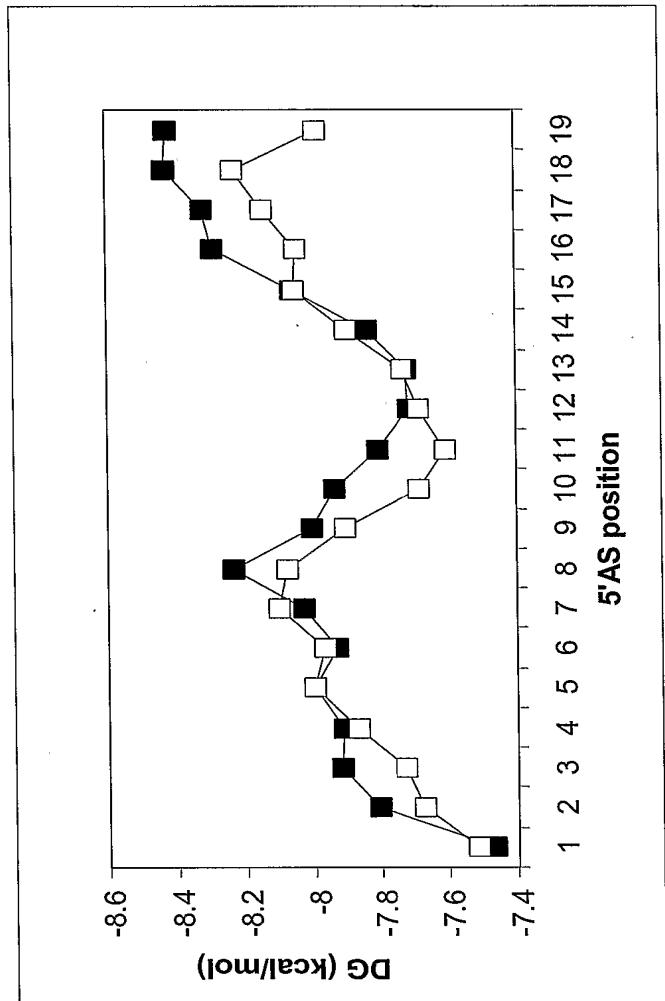
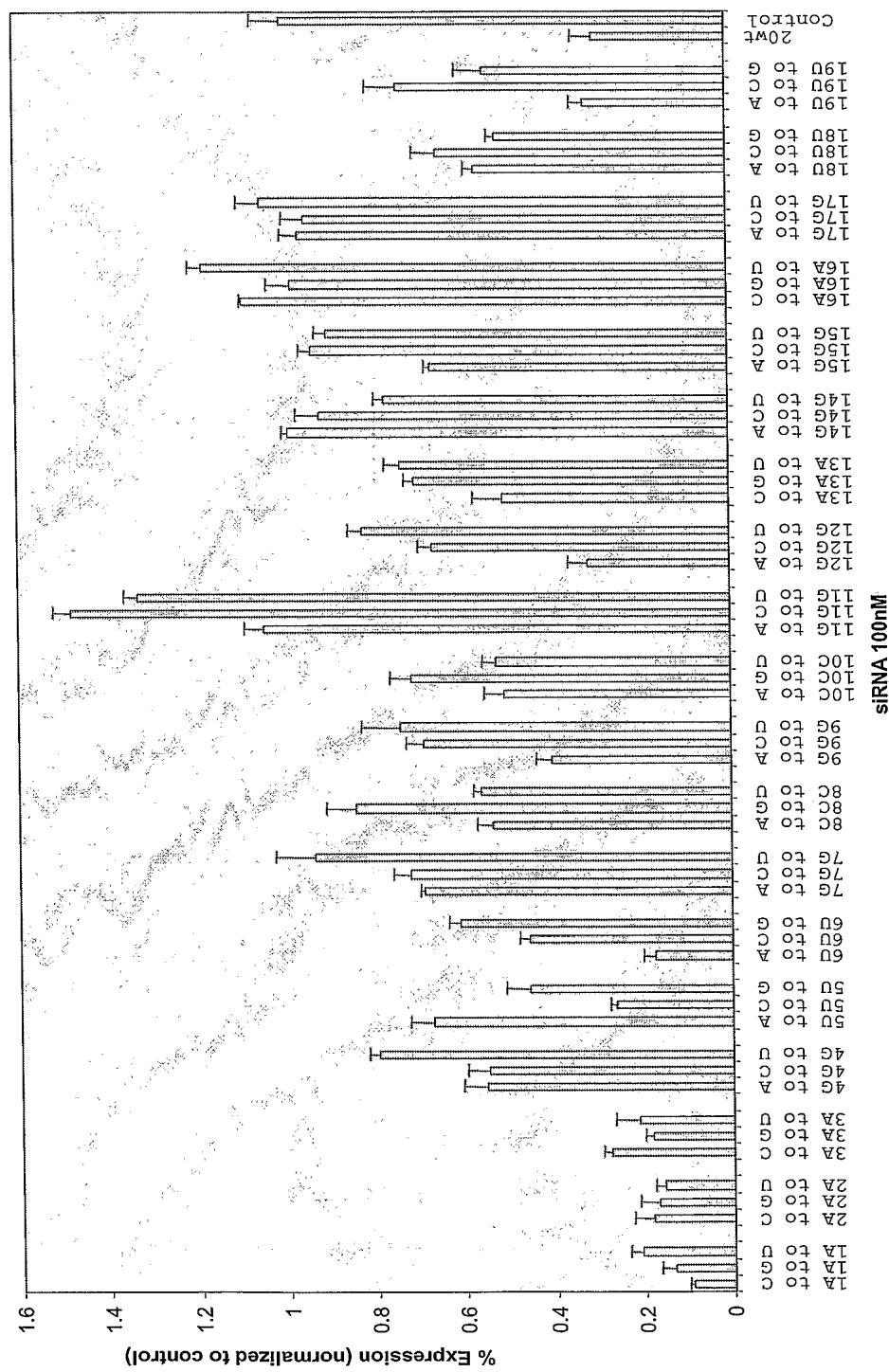
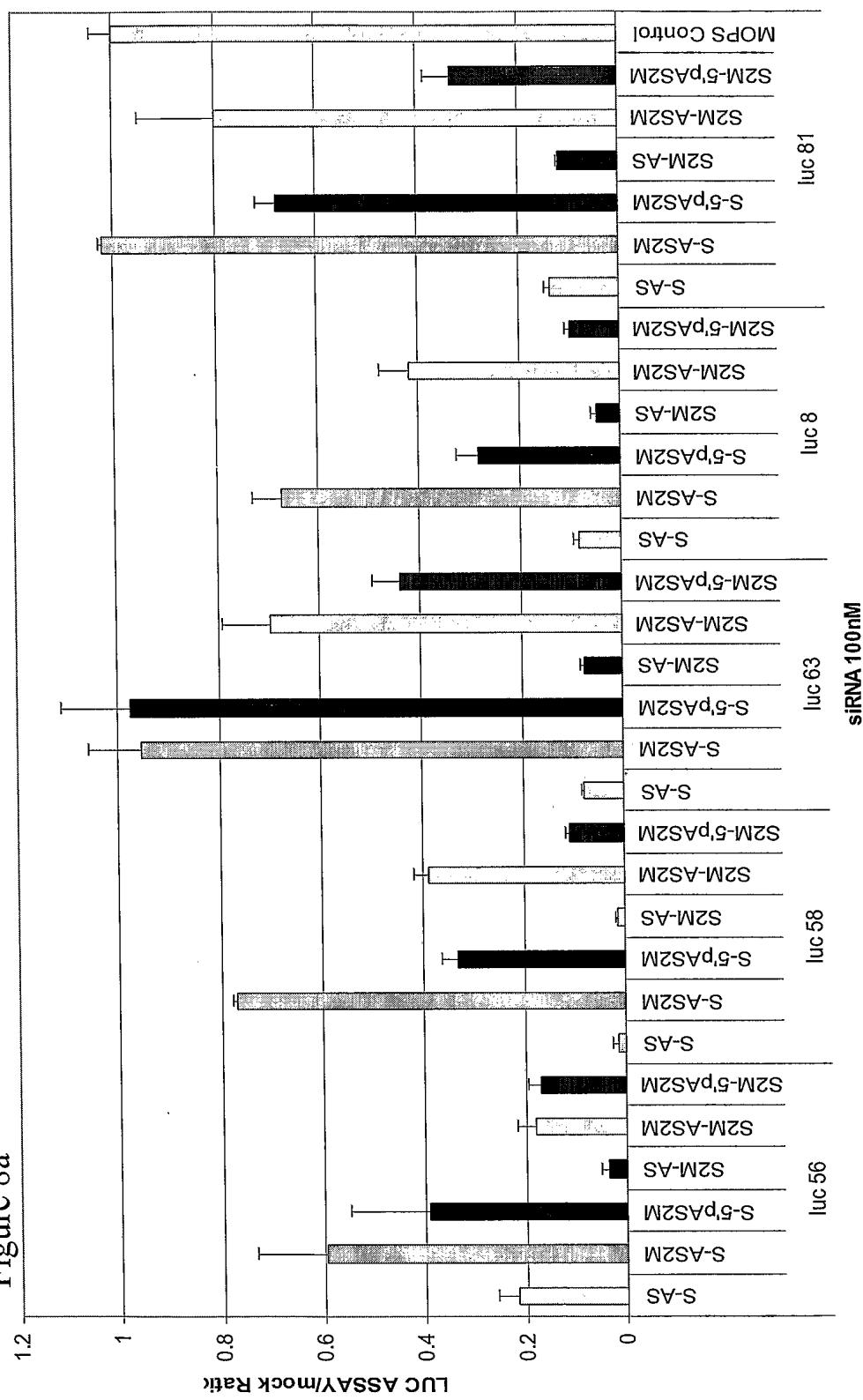


Figure 7



TARGET Screen Normalized LUC ASSAY 293 cells

Figure 8a



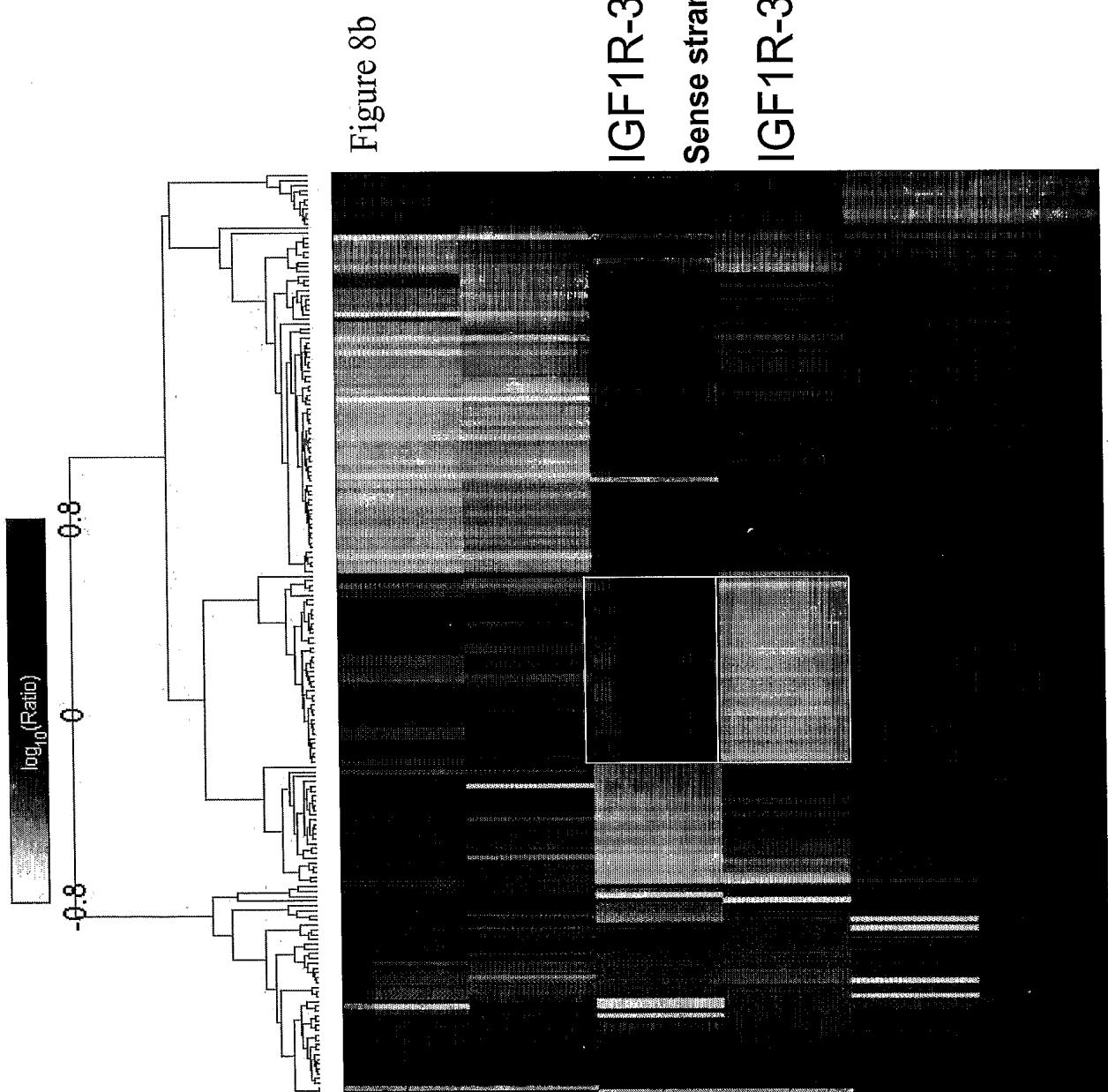


Figure 9

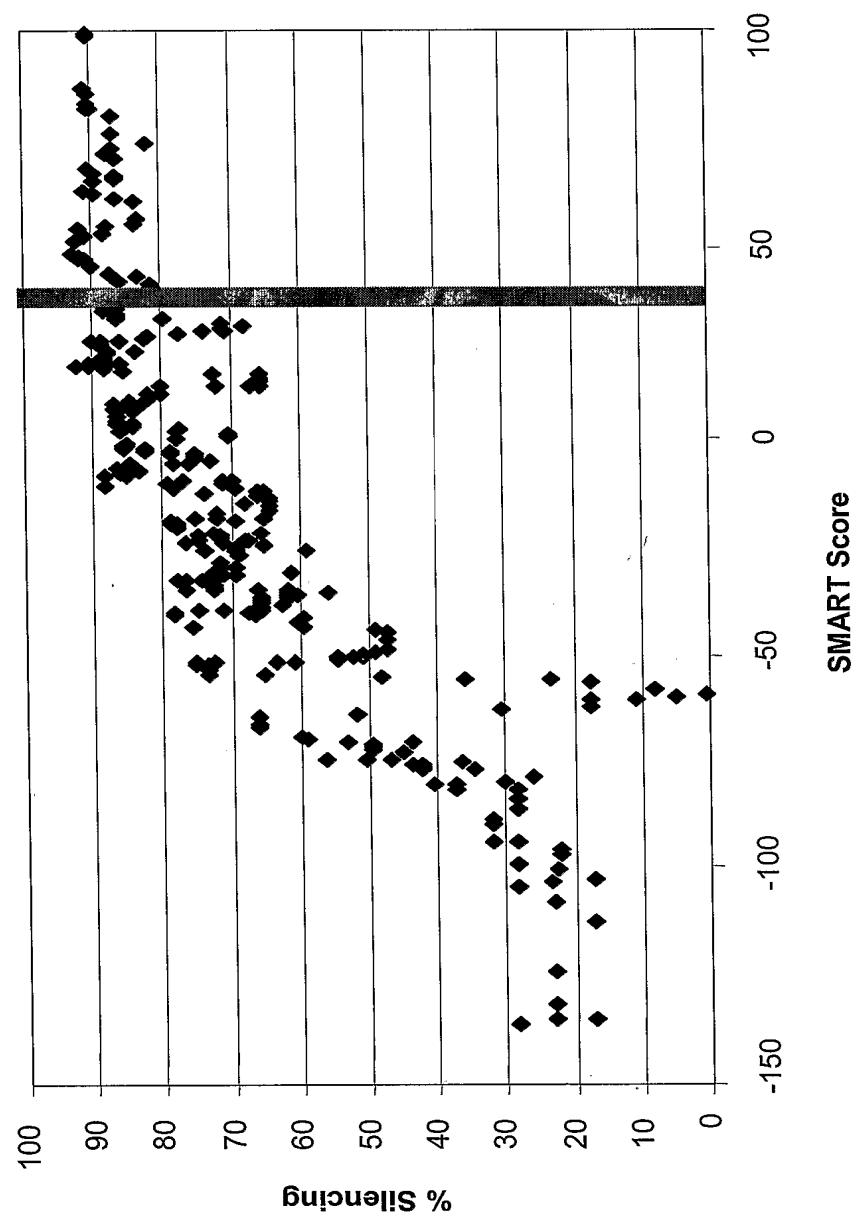


Figure 10a-f

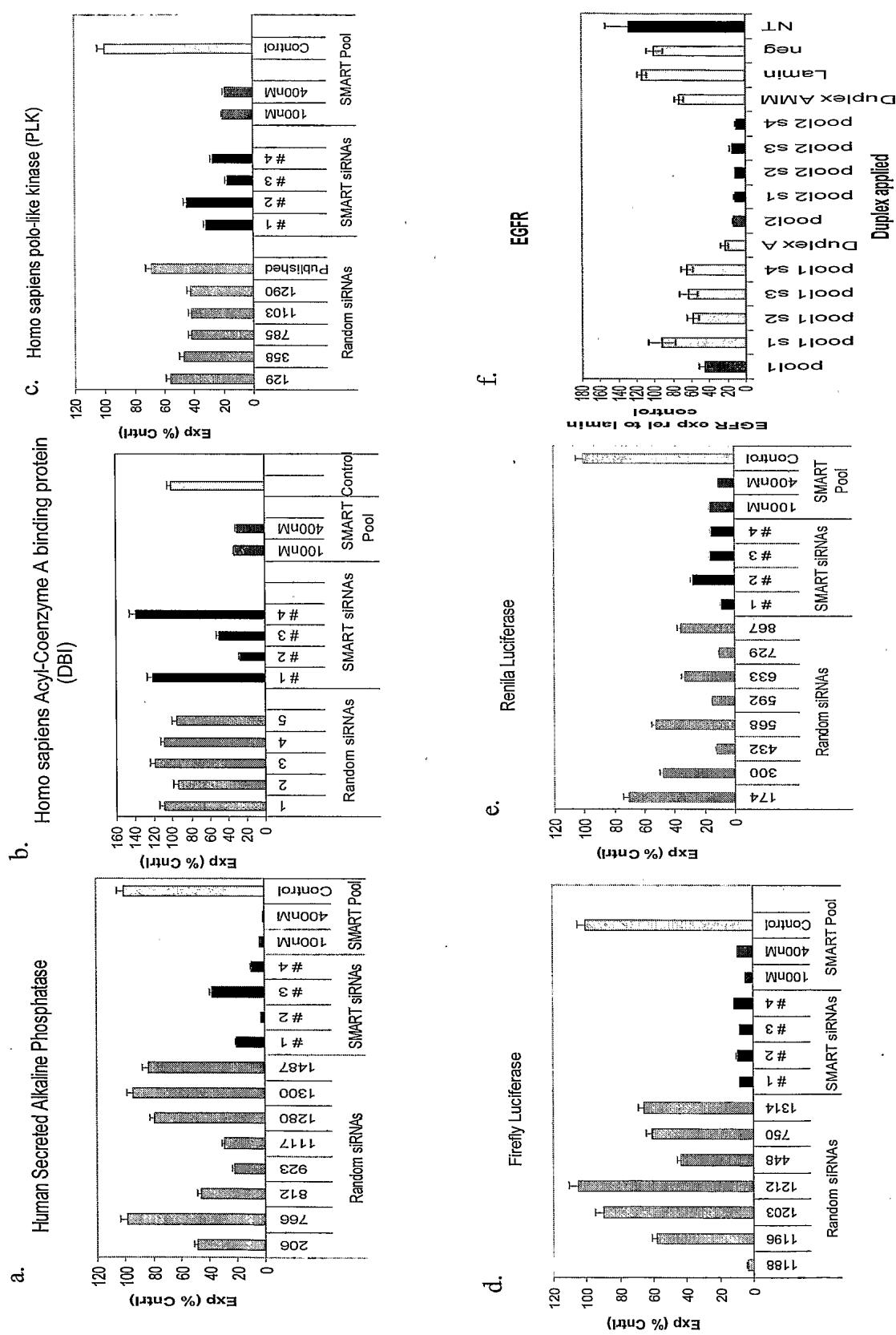
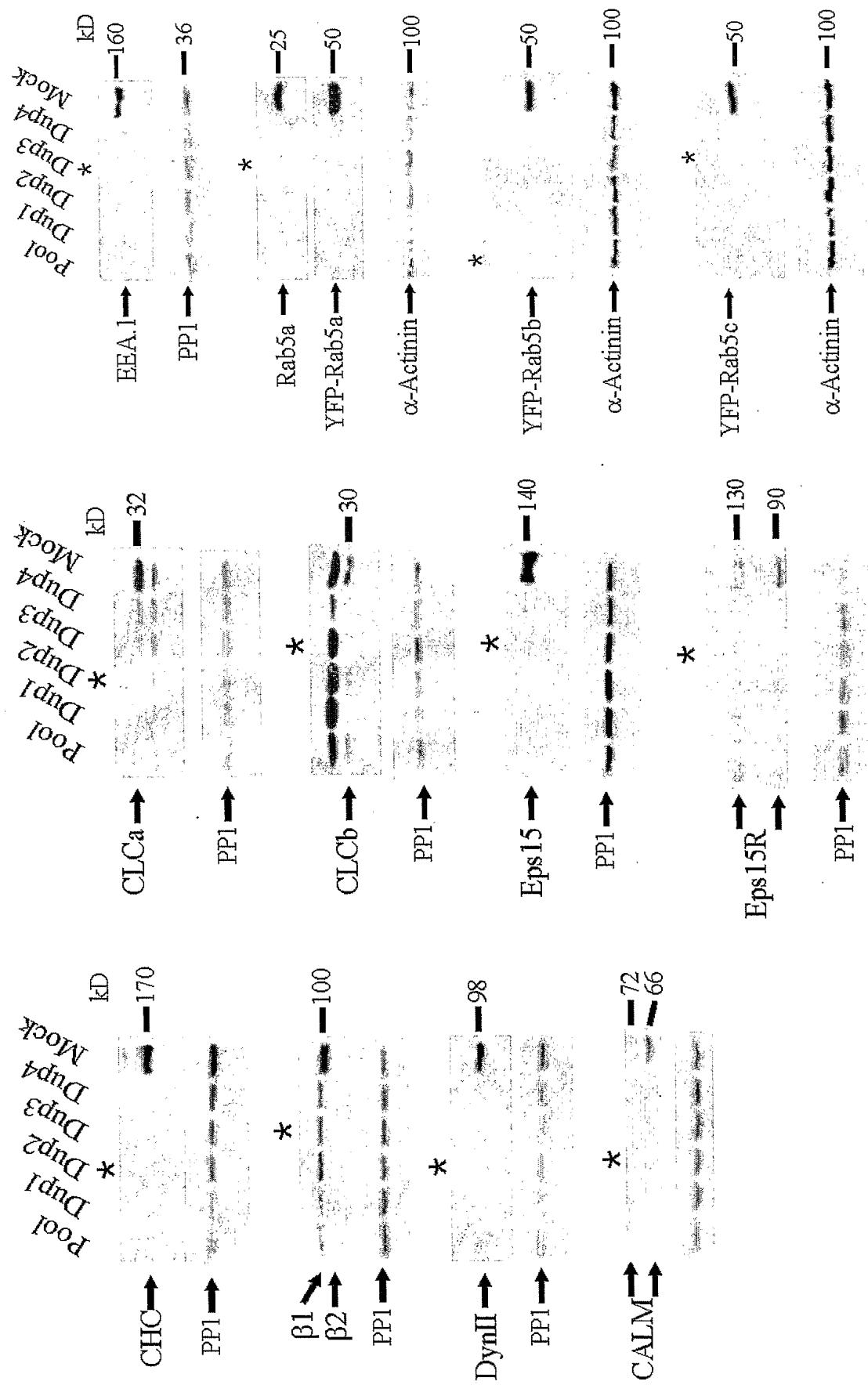


Figure 11



### Rational selection validation

Figure 12

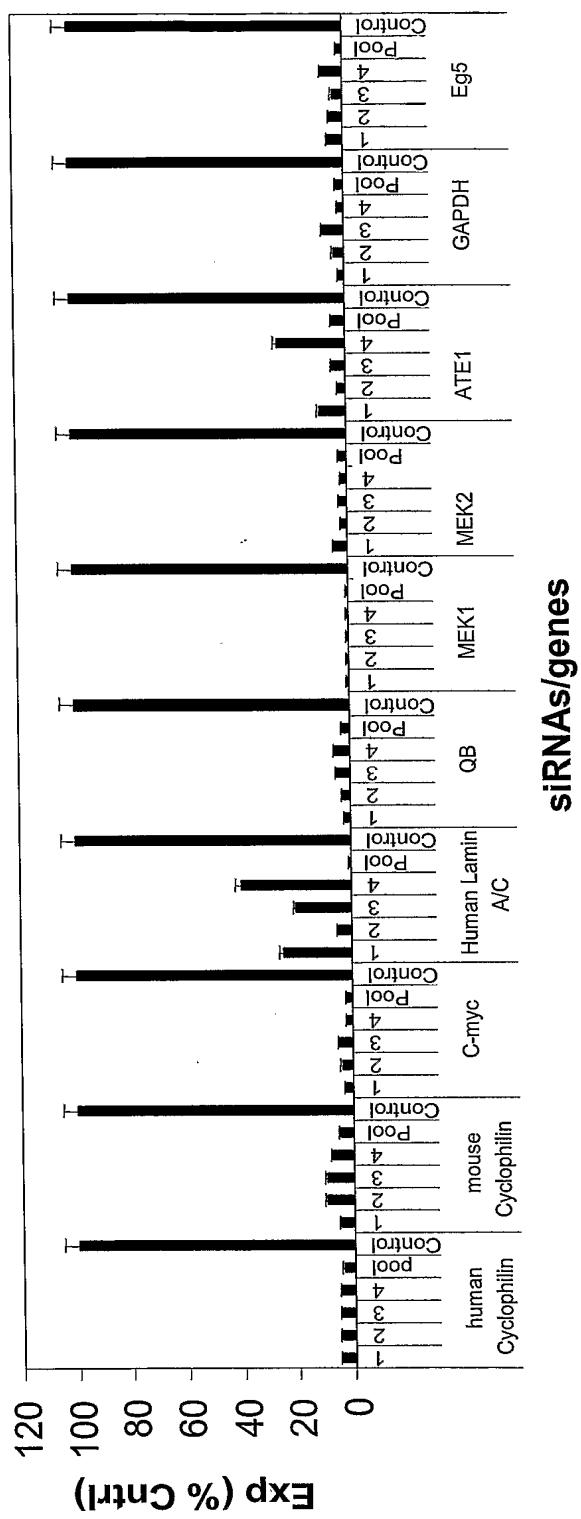
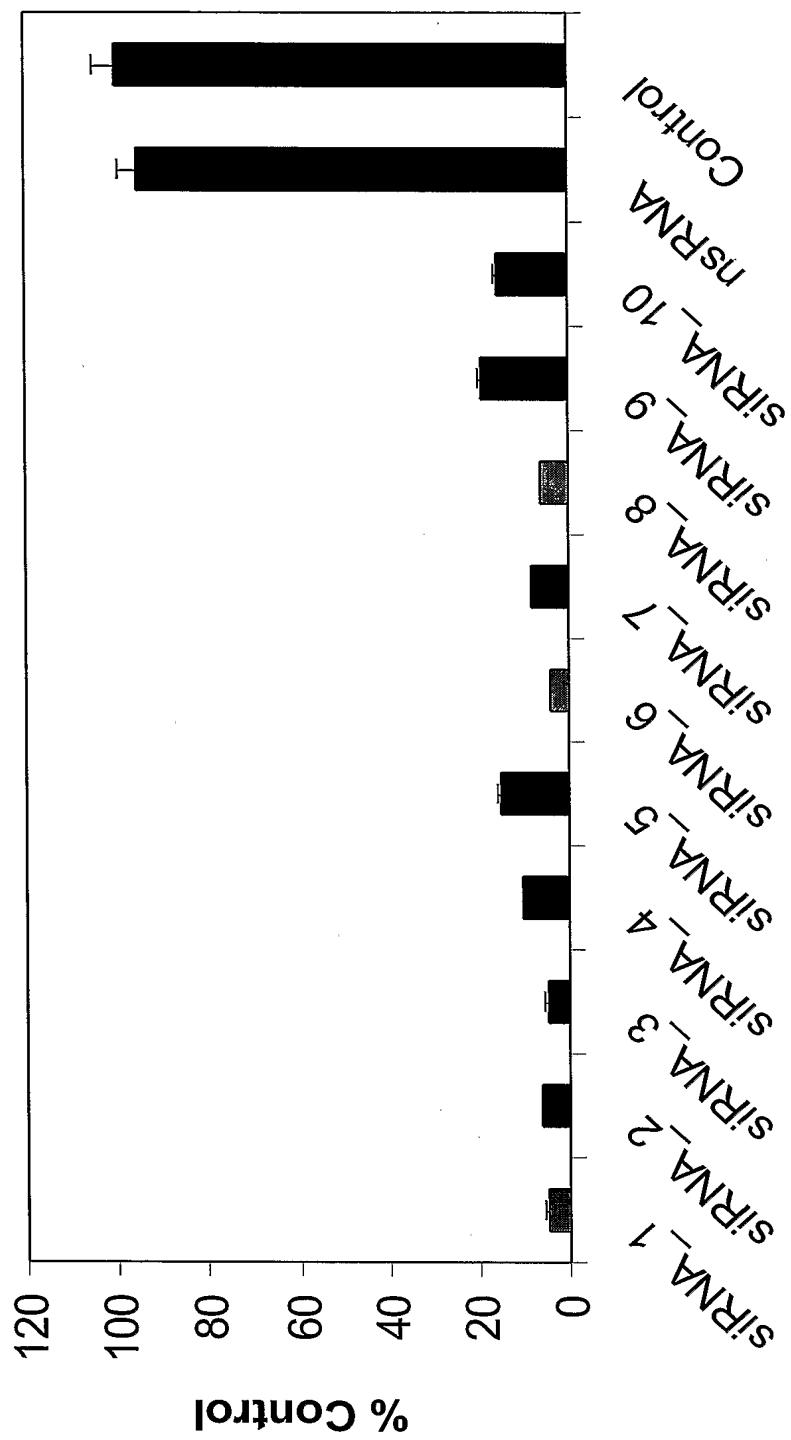


Figure 13 Sequences of top Bcl2

siRNA 1 GGGAGAUAGUGAUGAAAGUA  
siRNA 2 GAAGUACAUCCAUUAAG  
siRNA 3 GUACGACAACCGGGAGAU  
siRNA 4 AGAUAGUGAUGAAGUACAU  
siRNA 5 UGAAGACUCUGCUAGUUU  
siRNA 6 GCAUGGGCCUCUGUUUGA  
siRNA 7 UGCGGGCCUCUGUUUGAUUU  
siRNA 8 GAGAUAGUGAUGAAGUAC  
siRNA 9 GGAGAUAGUGAUGAAGUAC  
siRNA 10 GAAAGACUCUGCUAGUUUG

Figure 14

**Bcl-2 knockdown by 10 rationally designed siRNAs at 100 nM concentration**



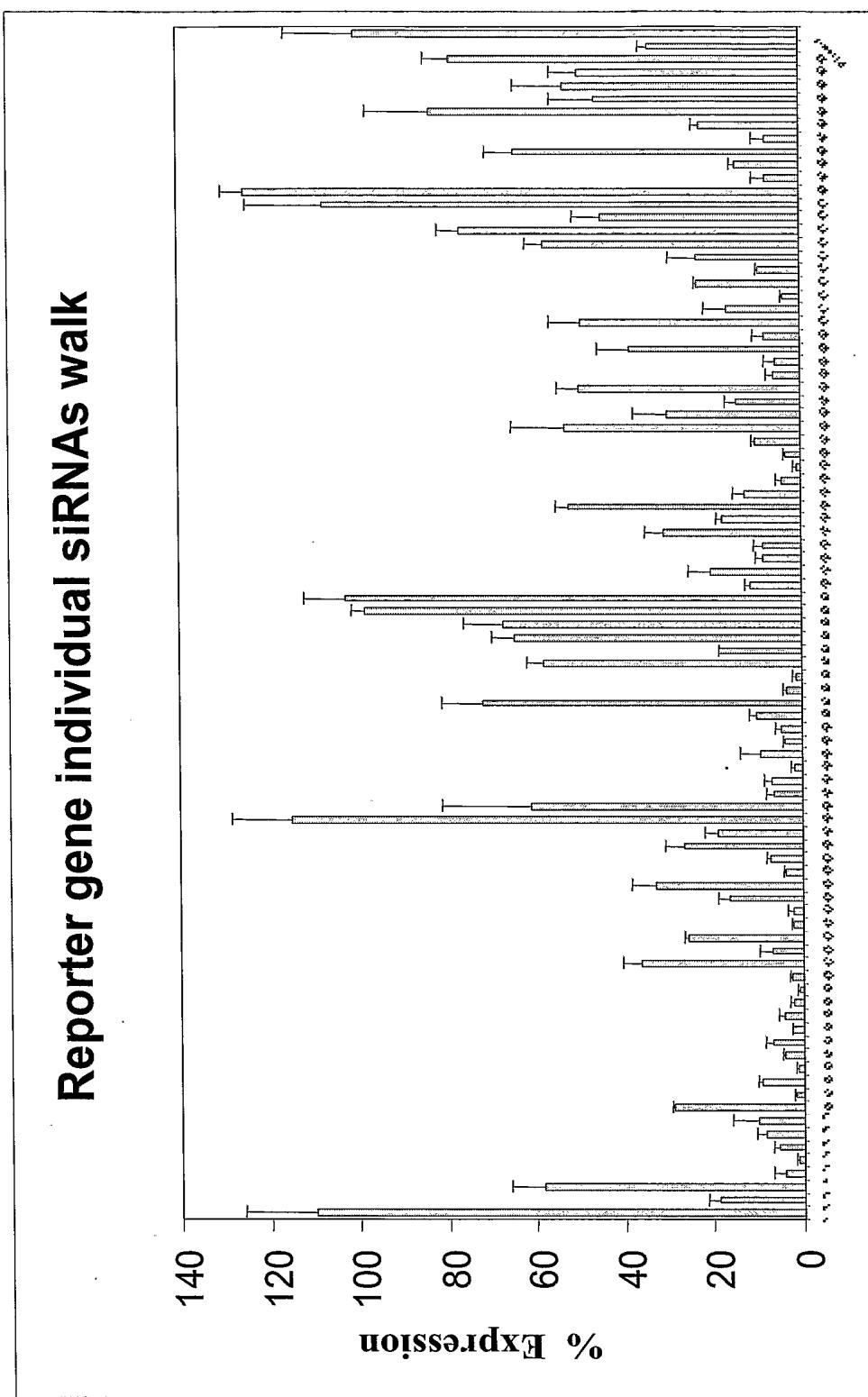


Figure 15

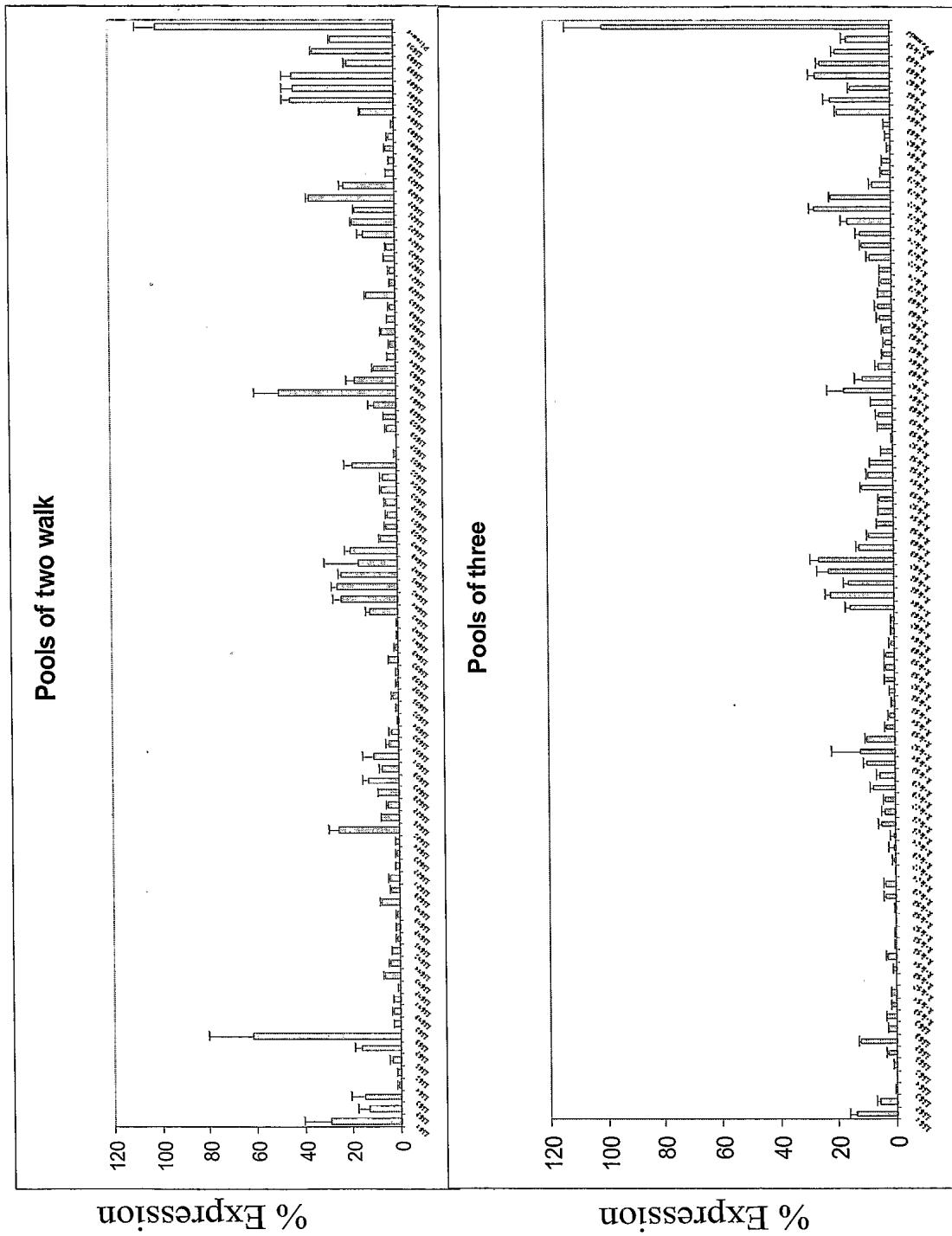


Figure 16

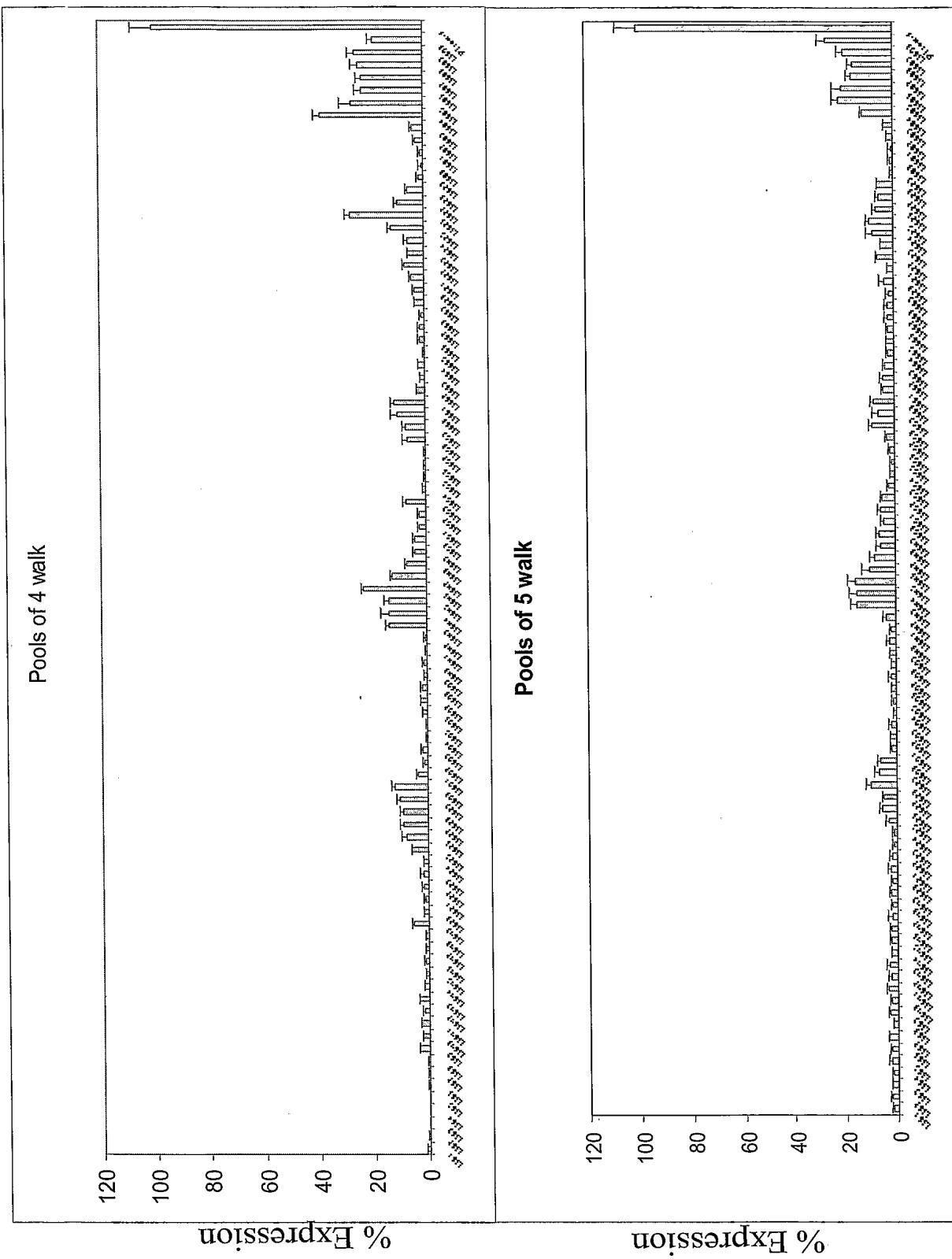


Figure 17

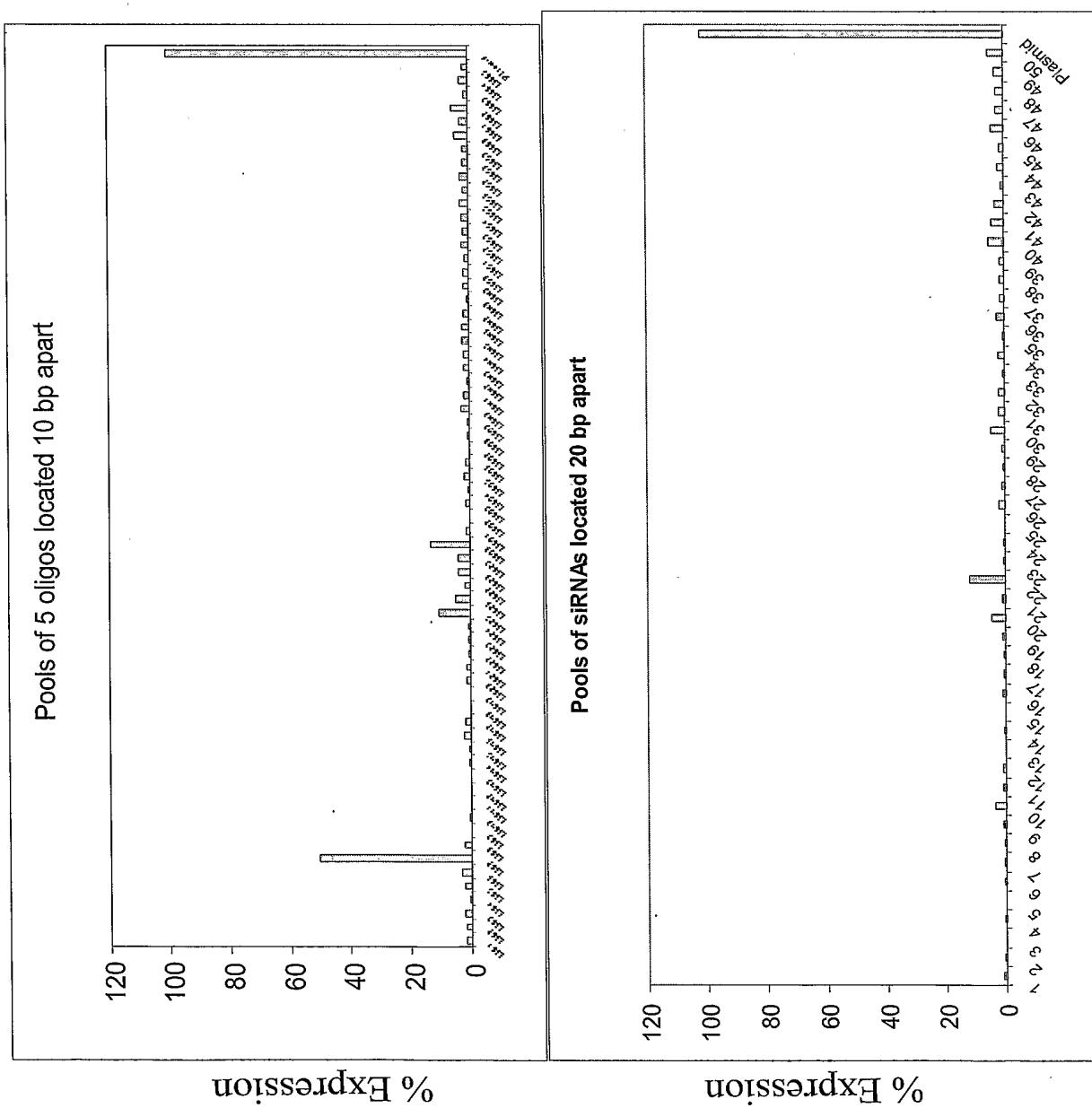


Figure 18

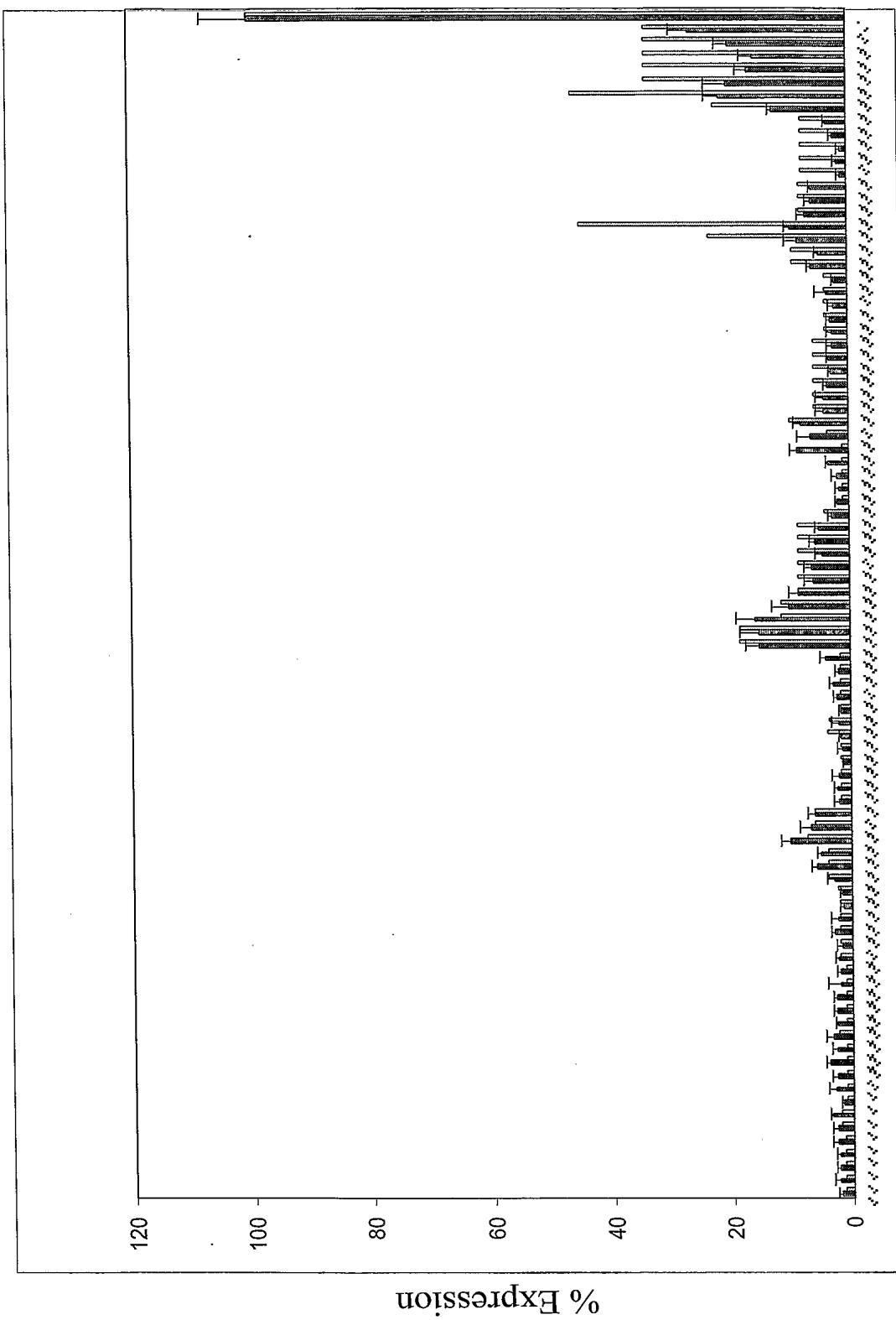


Figure 19

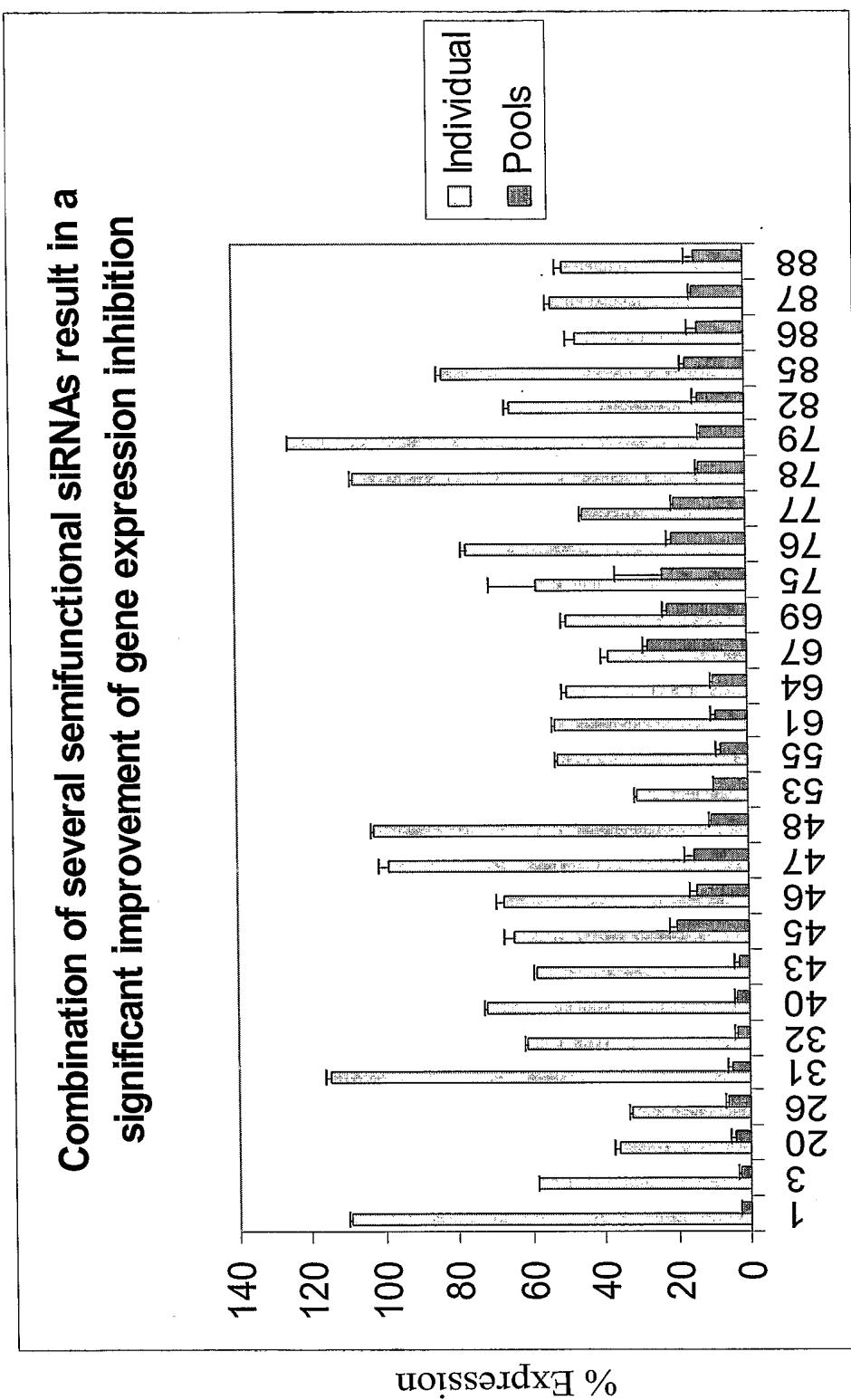


Figure 20

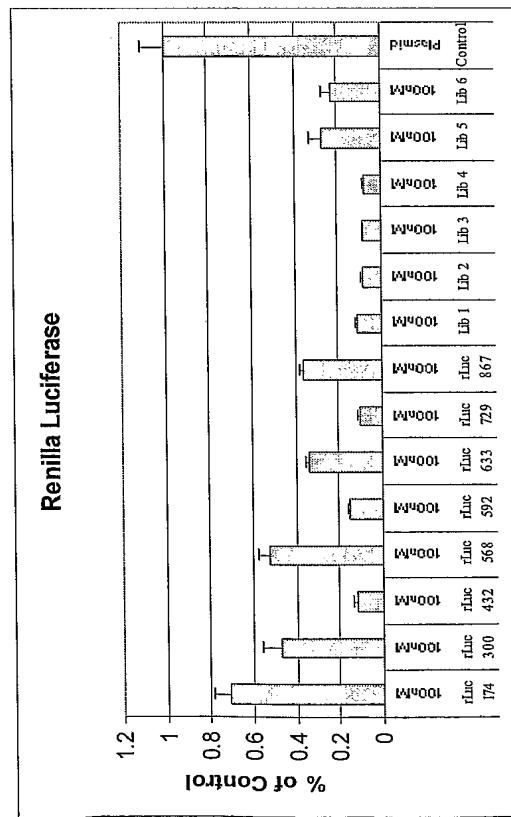
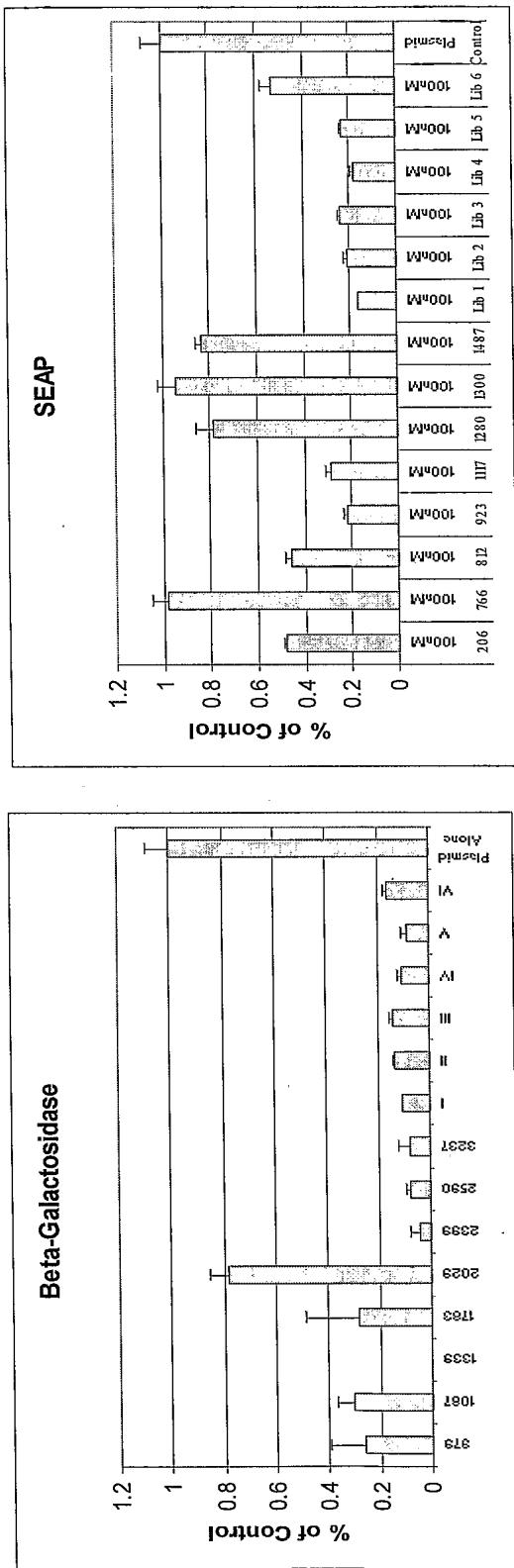


Figure 21

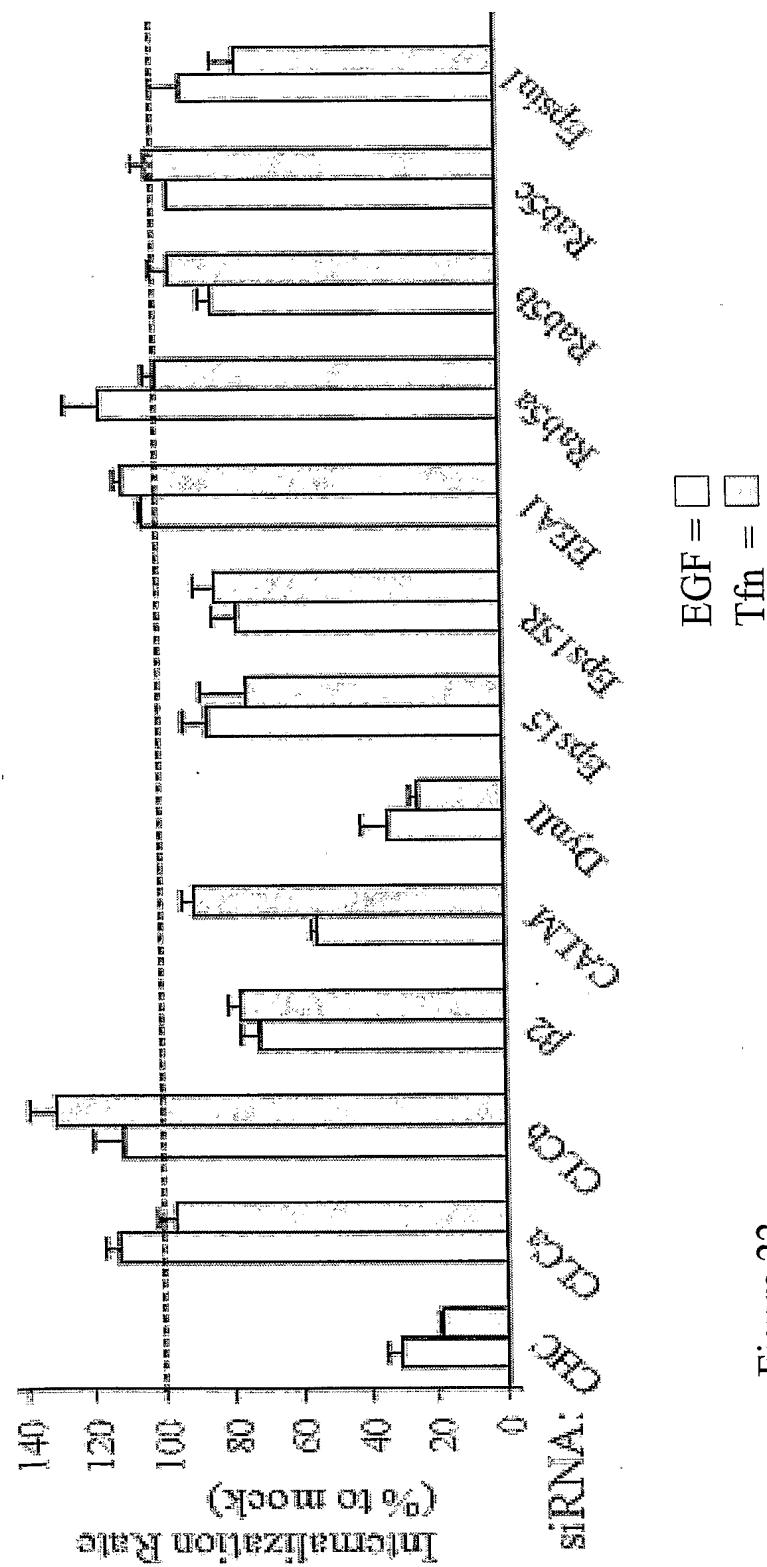


Figure 22

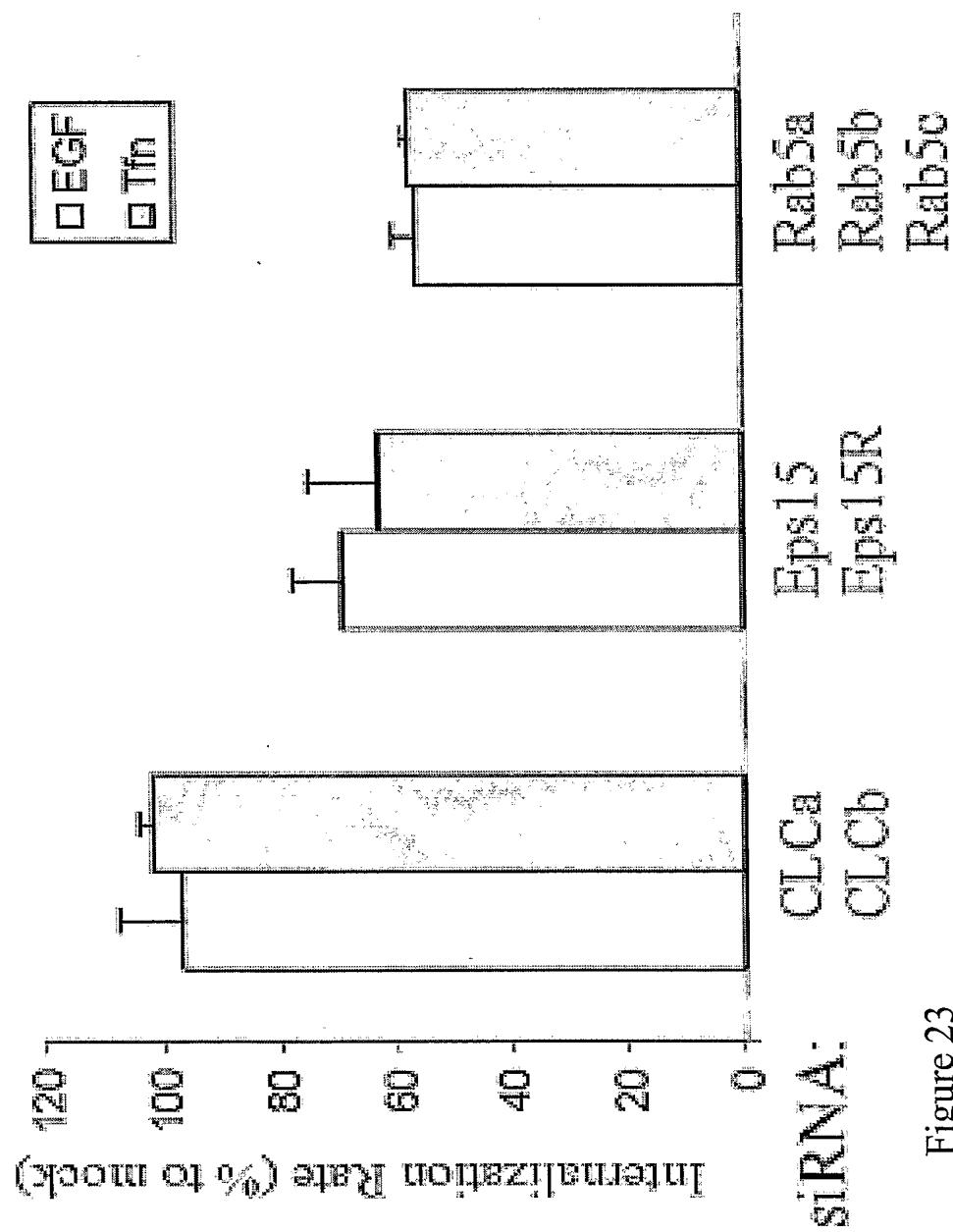


Figure 23

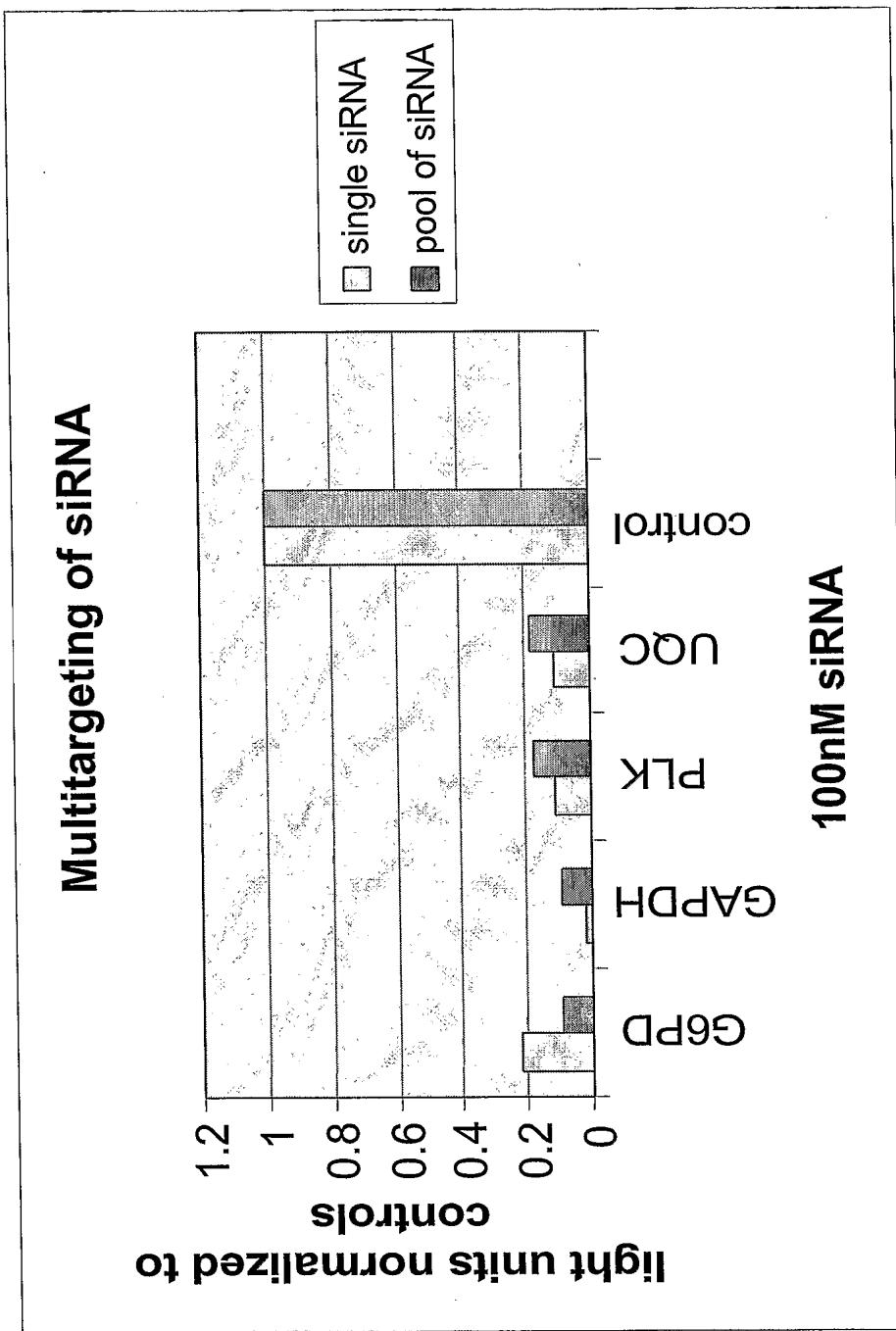
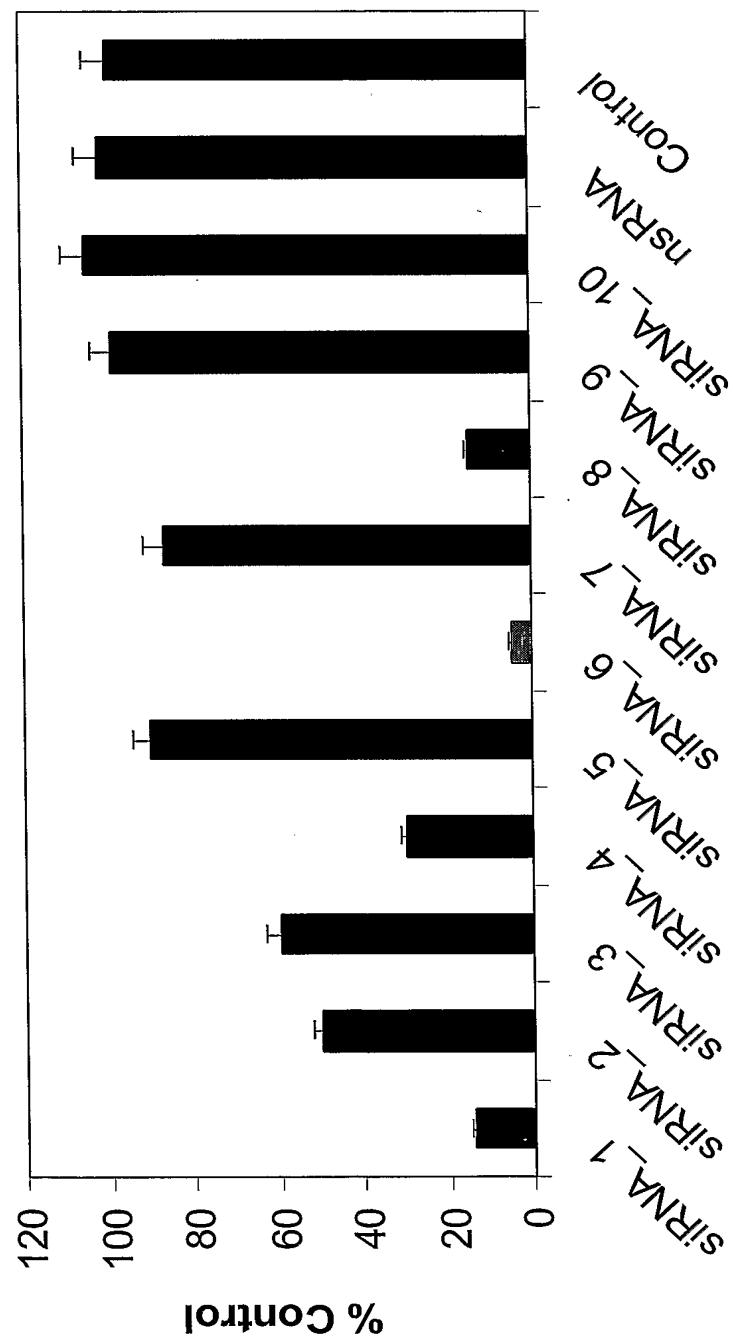


Figure 24

Figure 25

Bcl-2 knockdown by 10 rationally designed siRNAs at  
300 pM concentration



(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
3 June 2004 (03.06.2004)

PCT

(10) International Publication Number  
WO 2004/045543 A3

(51) International Patent Classification<sup>7</sup>: C12Q 1/68, C07H 21/00

(21) International Application Number: PCT/US2003/036787

(22) International Filing Date: 14 November 2003 (14.11.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
60/426,137 14 November 2002 (14.11.2002) US  
60/502,050 10 September 2003 (10.09.2003) US

(71) Applicant: DHARMACON, INC. [US/US]; 2650 Crescent Drive, Suite 100, Lafayette, CO 80026 (US).

(72) Inventors: ANASTASIA, Khvorova; 635 Walden #204, Denver, CO (US). ANGELA, Reynolds; 11445 Conifer Ridge Drive, Conifer, CO 80433 (US). DEVIN, Leake; 3050 Krameria Street, Denver, CO 80303 (US). WILLIAM, Marshall; 495 Mohawk Drive, Denver, CO (US).

(74) Agent: KALOW, David; Kalow & Springut LLP, 488 Madison Ave, 19th Floor, New York, NY 10022 (US).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VC, VN, YU, ZA, ZM, ZW.

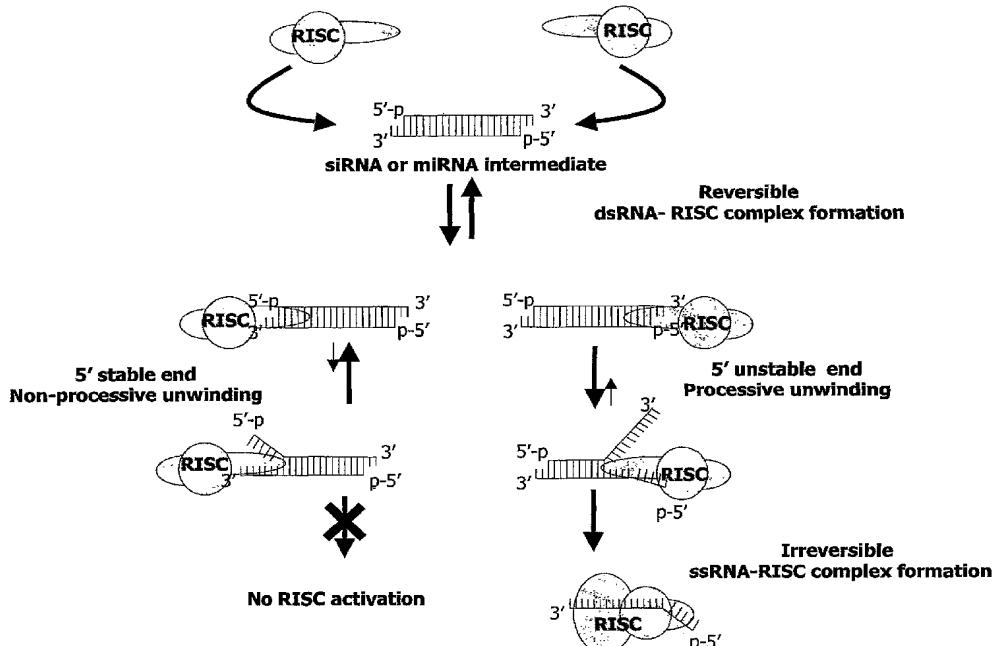
(84) Designated States (regional): ARIPO patent (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— with international search report

[Continued on next page]

(54) Title: FUNCTIONAL AND HYPERFUNCTIONAL SIRNA



WO 2004/045543 A3

(57) Abstract: Efficient sequence specific gene silencing is possible through the use of siRNA technology. By selecting particular siRNAs by rationale design, one can maximize the generation of an effective gene silencing reagent, as well as methods for silencing genes.



— before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(88) Date of publication of the international search report:

21 April 2005

# INTERNATIONAL SEARCH REPORT

International application No.

PCT/US03/36787

## A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) : C12Q 1/68; C07H 21/00  
US CL : 435/6; 536/24.5

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)  
U.S. : 435/6; 536/24.5

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)  
Please See Continuation Sheet

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category * | Citation of document, with indication, where appropriate, of the relevant passages   | Relevant to claim No. |
|------------|--|-----------------------|
| A          | US 2002/0150945 A1 (FINNEY et al) 17 October 2002 (17.1 O.2002), see page 27-28.   | 1-19                  |
| A          | KASIF et al. A computational framework for optimal masking in the synthesis of oligonucleotide microarrays.<br>Nucleic Acids Research. 2002, Vol. 30, No. 20, full text version of article.  | 1-19                  |
| A          | AMARZGUIQUI et al. Secondary structure prediction and <i>in vitro</i> accessibility of mRNA as tools in the selection of target sites for ribozymes.<br>Nucleic acids Research, 2000, Vol. 28, No. 21, pages 4113-4124, full text version. | 1-19                  |

Further documents are listed in the continuation of Box C.

See patent family annex.

|   |     |  |
|---|-----|--|
| * Special categories of cited documents:  | "T" | later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention  |
| "A" document defining the general state of the art which is not considered to be of particular relevance  | "X" | document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone   |
| "E" earlier application or patent published on or after the international filing date   | "Y" | document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art |
| "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) | "&" | document member of the same patent family  |
| "O" document referring to an oral disclosure, use, exhibition or other means  |     |  |
| "P" document published prior to the international filing date but later than the priority date claimed  |     |  |

Date of the actual completion of the international search

19 January 2005 (19.01.2005)

Date of mailing of the international search report

25 FEB 2005

Name and mailing address of the ISA/US

Mail Stop PCT, Attn: ISA/US  
Commissioner for Patents  
P.O. Box 1450  
Alexandria, Virginia 22313-1450

Facsimile No. (703) 305-3230

Authorized officer

Janet L. Epps-Ford, Ph.D.

Telephone No. 571-272-0547

**INTERNATIONAL SEARCH REPORT**

PCT/US03/36787

**Continuation of B. FIELDS SEARCHED Item 3:**

CAplus, Medline, Biosis, USPatfull, Derwent, JPO, EPO

search terms: (SIRNA OR RNAI OR DSRNA) and (OPTIMIZATION OR OPTIMIZE OR OPTIMAL) and algorithm